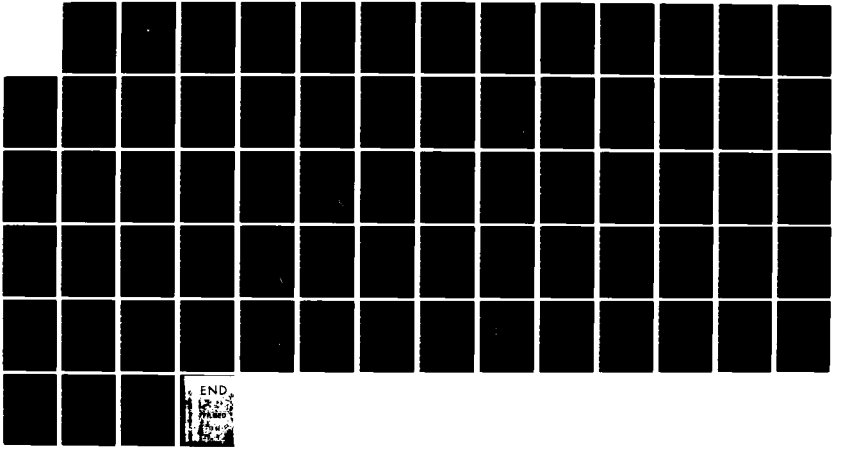
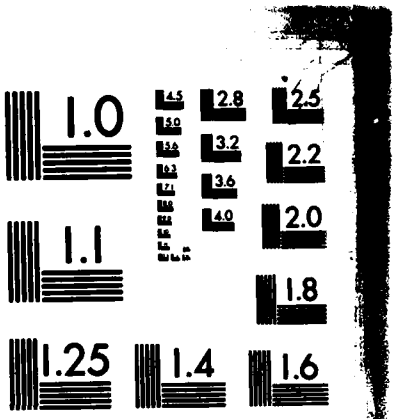
AD-A133 349 DATA ANALYSIS AND MODELING OF ARCTIC SEA ICE SUBSURFACE ROUGHNESS(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA D P GAVER ET AL. NOV 82 NPS55-82-031 1/1

UNCLASSIFIED F/G 8/12 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NPS55-82-031

NAVAL POSTGRADUATE SCHOOL

Monterey, California

DATA ANALYSIS AND MODELING OF ARCTIC SEA ICE

SUBSURFACE ROUGHNESS

by

D. P. Gaver

P. A. Jacobs

November 1982

Prepared for:
Chief of Naval Research
Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral J. J. Ekelund
Superintendent

David A. Schrady
Provost

This work was supported in part by the Naval Postgraduate School Foundation Research Program and by the Office of Naval Research, Arlington, VA.

P. A. Jacobs, Associate Professor
Department of Operations Research

D. P. Gaver, Professor
Department of Operations Research

Reviewed by:

K. T. Marshall, Chairman
Department of Operations Research

Released by:

William M. Tolles
William M. Tolles
Dean of Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55-82-031	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DATA ANALYSIS AND MODELING OF ARCTIC SEA ICE SUBSURFACE ROUGHNESS		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) D. P. Gaver P. A. Jacobs		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61152N; RR000-01-10 N0001482WR20043
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE November 1982 13. NUMBER OF PAGES 67
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Arctic ice
ice keels
ice distribution
statistical distributions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Statistical data analysis and models are used to characterize and summarize the roughness of the underside of sea ice in the Arctic. Keel spacings and depths are modeled by sculptured exponentials, and by gamma distributions. The data studied was obtained by upward looking sonar on the submarine GURNARD during April, 1976, in the Beaufort Sea. The models and methods should be more widely applicable.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DATA ANALYSIS AND MODELING OF ARCTIC SEA ICE

SUBSURFACE ROUGHNESS

D. P. Gaver

P. A. Jacobs

Department of Operations Research
Naval Postgraduate School
Monterey, California 93940

Executive Summary

Arctic sea-ice behavior, particularly roughness and ridging (keeling) patterns above and below the surface, is of scientific interest to oceanographers and geologists. It is also of potential interest to those conducting military operations in the Arctic, and to those exploring for petroleum and other minerals. In particular, military operations involving submarines are facilitated by access to the surface from below, so indications of the distribution of "leads" or "polynyas", or of relatively thin ice regions, are of interest.

This paper investigates the statistical distribution of relatively deep keels occurring beneath the ice. Such keels may provide obstacles to under-ice vehicles; detached, they may be the agents by which gouging of the bottom occurs. Such gouges threaten pipelines or cables.

The present work is based upon data furnished by Dr. Peter Wadhams of Scott Polar Research Institute, Cambridge, England. It was originally obtained by upward-looking sonar aboard the submarine U.S.S. GURNARD during April, 1976, for a route beneath the Beaufort Sea ice.

The methods utilized are those of exploratory data analysis and of fitting apparently suitable statistical distributions (probability models) to the distances between successive deep (> 30 ft.) keels, and to the depths of the keels identified. The data suggest that an exponential-like, but not

precisely exponential, model may well represent the data: the simple exponential "model" favored by others previously tends to underestimate the distances between keels, and, perhaps more importantly, to underestimate the extreme keel depths. This is of interest, for deep projections are most likely to impede vehicle progress, and to cause gouging.

The present numerical summaries are of a pilot study. The methodology used has more and more widely applicable elements: the "sculptured exponential" distributions utilized here may well be of service for summarizing other data involving the environment.

DATA ANALYSIS AND MODELING OF ARCTIC SEA ICE SUBSURFACE ROUGHNESS

D. P. Gaver
P. A. Jacobs

Department of Operations Research
Naval Postgraduate School
Monterey, California 93940

1. Introduction

The spatial pattern of the sea ice cover in the Arctic has been of considerable scientific interest to geophysicists and oceanographers for some time. Its presence importantly affects the environment for naval and other military operations, and for oil and mineral exploration. In particular, naval submarine operations are influenced by the existence of deep downward projections ("ice keels") from the surface canopy, by acoustic wave reflections from the underside of that canopy, and by the apparently random incidence of essentially open regions in the ice pack ("leads" or "polynyas") that permit access to the surface from below.

In this paper statistical methods are used to characterize and summarize features of the Arctic ice pack related to those mentioned above. The analysis is based on a particular set of data furnished by Dr. Peter Wadhams of Scott Polar Research Institute, Cambridge, England, to whom we are grateful. A previous analysis of these data has been reported by Wadhams and Horne [1980], hereafter abbreviated WH. While the approaches of earlier investigators have lead to simple one-parameter exponential distributions as summaries of data describing (a) spatial intervals between keel occurrences, and also (b) keel depths, our definitions and data analysis suggest that both keel spacings and keel depths are longer-tailed than the exponential. We further suggest simple parametric forms to summarize the observed statistical behavior.

2. The Data

The data we analyze were obtained by upward-looking sonar aboard the submarine U.S.S. GURNARD during the period April 7-10, 1976, from beneath the Beaufort Sea ice canopy. The route followed by the GURNARD was from a point north of Barter Island (just over 70^{o} N.) to 75-76^{o} N, thence southeasterly to a point 72-73^{o} N, and finally westerly to a point northeast of Pt. Barrow. For a detailed map see WH. The data -- ice drafts, measured from below the ice to the surface -- were taken over a 1400 km. transect length. Data tapes were initially cleaned and processed at the Arctic Submarine Laboratory, Naval Undersea Center, San Diego; they were later further processed at Scott Polar Research Institute, and observations which were taken at intervals of 1.3 - 1.5 m. were referred by interpolation to a nominal 1.0 m. spacing. Furthermore, the data file was split into sections, each of which make up about 50 km of data. There were 27 such sections, with a gap appearing between two of them. More detail is available in WH.

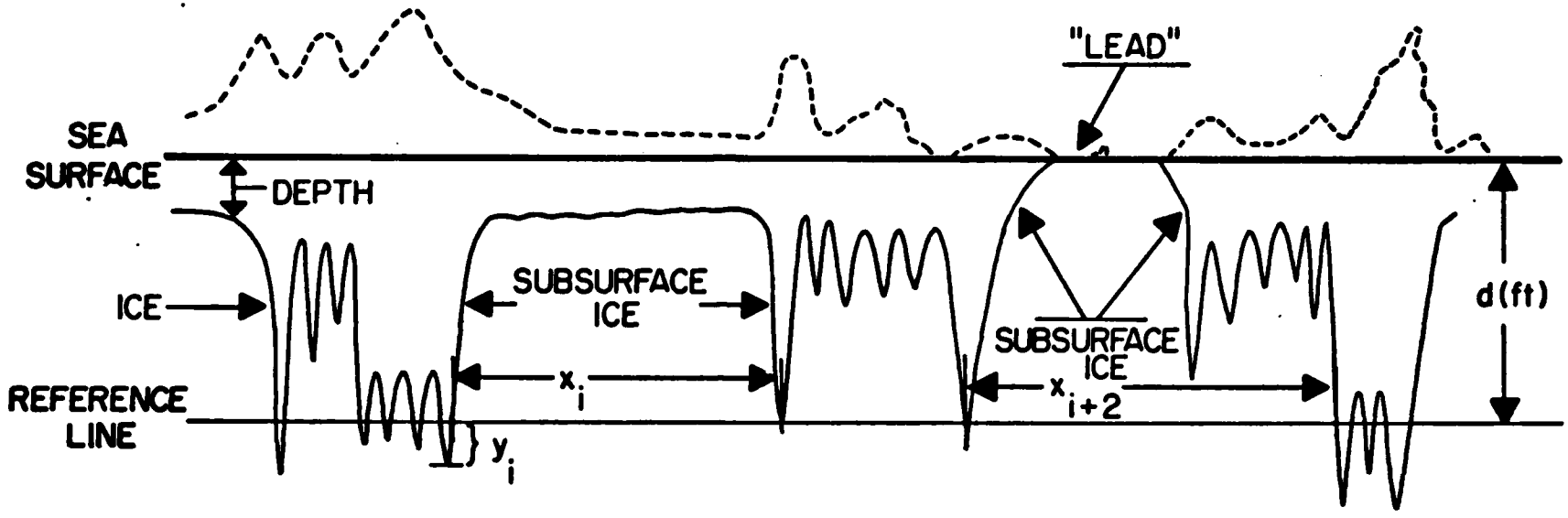
Certainly the data set referred to can provide considerable information concerning the underside of Arctic ice. However, there are recognizable limitations in the inferences that may be well-justified from even a sophisticated analysis of these particular measurements. For instance, the data were obtained during a relatively short period of time in one year, so there is no opportunity to assess month-to-month or season-to-season variability. In order to obtain more information, more data must be subjected to analysis. Our purpose here is to suggest methods of analysis that may be useful when such data become available.

3. Data on Keel Spacings and Keel Depths: Definitions

The raw data on ice drafts were transformed into data on keel spacings and magnitudes by the simple expedient of constructing an imaginary line, L , at a constant depth d(feet) below sea surface, and then measuring distances (spacings) between successive up- and down-crossings of L , denoted generically by x , and the maximum depth (keel depth, relative to d) achieved between a down-crossing and the first subsequent up-crossings, denoted by y . Figure 1 should clarify this definition, which differs somewhat from that of WH: it permits the occurrence of more small spacings than does theirs.

Data on spacings and keels were initially obtained for three levels: d = 30 (feet), 40, 50. These depths are apparently of interest from a submarine operational view point, but are too deep to be of great interest to acousticians. Further analysis of crossings at smaller depths is in progress.

ARCTIC SEA-ICE DATA APPEARANCE



x_i = DISTANCE BETWEEN UP-CROSSING AND SUBSEQUENT DOWN CROSSING OF REFERENCE LEVEL (PARALLEL TO SURFACE) AT DEPTH d (e.g. 30', 40', 50').

y_i = KEEL DEPTH, WITH RESPECT TO REFERENCE LEVEL, DEPTH d.

Figure 1.

4. Exploratory - Descriptive Analyses of Spacings

The initial step in the analysis of spacings was to select a data segment running from latitude 71.140, longitude 144.225 to latitude 74.328, longitude 144.378 and to graphically display the spacings in various ways to see if any apparent patterns emerged. We discuss here spacings at a depth of 30 ft.

(a) Serial plot. This is a plot of x_i vs i , where i indicates the order in which the spacing occurred along the track; i = 1 refers to the first spacing, i = 2 to the second encountered, etc. Such a plot appears in Figure 2 . There is some visual evidence that fewer long spacings occur among the first 200 or so (out of more than 600), hence that the series may be somewhat non-stationary.

(b) Serial histograms. The data were segregated or binned in groups of size 73 (convenient fraction of total number) in the order of their occurrence, and each group was histogrammed; see Figure 3 . This presentation reveals the exponential-like positive skewness (J-shapedness) of the data, and also suggests that long spacings tend to occur late along this segment of the track.

(c) Serial boxplots. The groups of 73 were next box-plotted; see Tukey [1977], and the means joined by an eye-guiding line. Once again the picture indicates that the longest spacings tend to occur later along the track. There is a slight upward trend noticeable in the mean line that is probably attributable to the influence of the largest spacings in each group; a similar plot connecting medians would not likely show much trend; Figure 4.

(d) Histogram of log (spacings), all data. If data are highly skewed, some form of symmetrizing transformation is often useful; see Tukey [1977] McNeil [1977]. The log transformation tends to

symmetrize extreme positive skewness (as in our histograms or in an exponentially distributed sample); it may also reveal hidden patterns; see McNeil [1977] p. 11. Figures 5 and 6 show the effect of such a transformation on the histogram of the 30 ft. spacings data: two roughly symmetrical but pronounced concentrations of values, or "bumps", become evident. The lower bump can be identified with very closely-spaced up and down crossings occurring across the bottom of a large ice structure (keel); the present way of identifying keels allows these "pseudo keels", which are usually defined by relatively shallow gouges, to appear, while the approach of WH. suppresses them. The upper bump, made up of about half the data at d = 30 ft., represents the genuine keel spacings believed to be of operational significance. The lower bump describes high-frequency keel occurrences likely to be of interest to acousticians, but this subset of data is also perturbed by measuring instrument noise; these data should first be smoothed to minimize the noise contribution. At d = 30 ft. the split between the lower and upper bumps identified occurs at nearly the median of the data, or at about 70 m.

On the basis of these last observations the analysis to follow will focus on attempts to summarize parametrically the distribution of data associated with the upper bump identified in the log-plot. These data appear to be comparable to the spacing data discussed by WH. and others. They appear also to have the most significance to those concerned with subsurface vehicular (submarine) operations.

More precisely, the first 200 observations of the series are dropped. The median of the remaining data is calculated and those data points less than the median are dropped. Finally, the median is subtracted from the remaining data points. Results of the analysis of this upper half of the spacings data will follow.

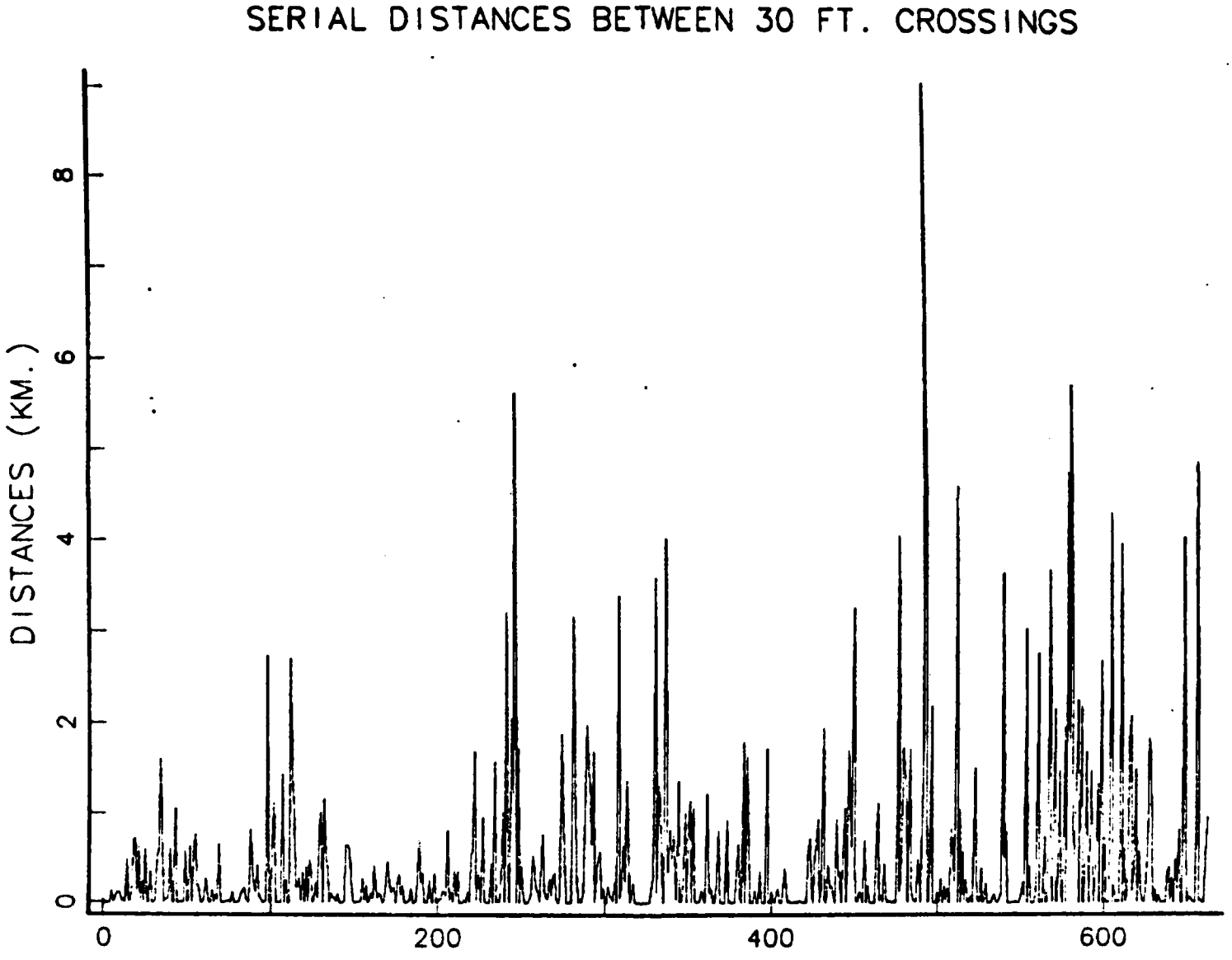

SERIAL DISTANCES BETWEEN 30 FT. CROSSINGS
DISTANCES (KM.)
0
2
4
6
8
0
200
400
600

FIGURE 2

SERIAL HISTOGRAMS OF SPACINGS (D=30 FT.)

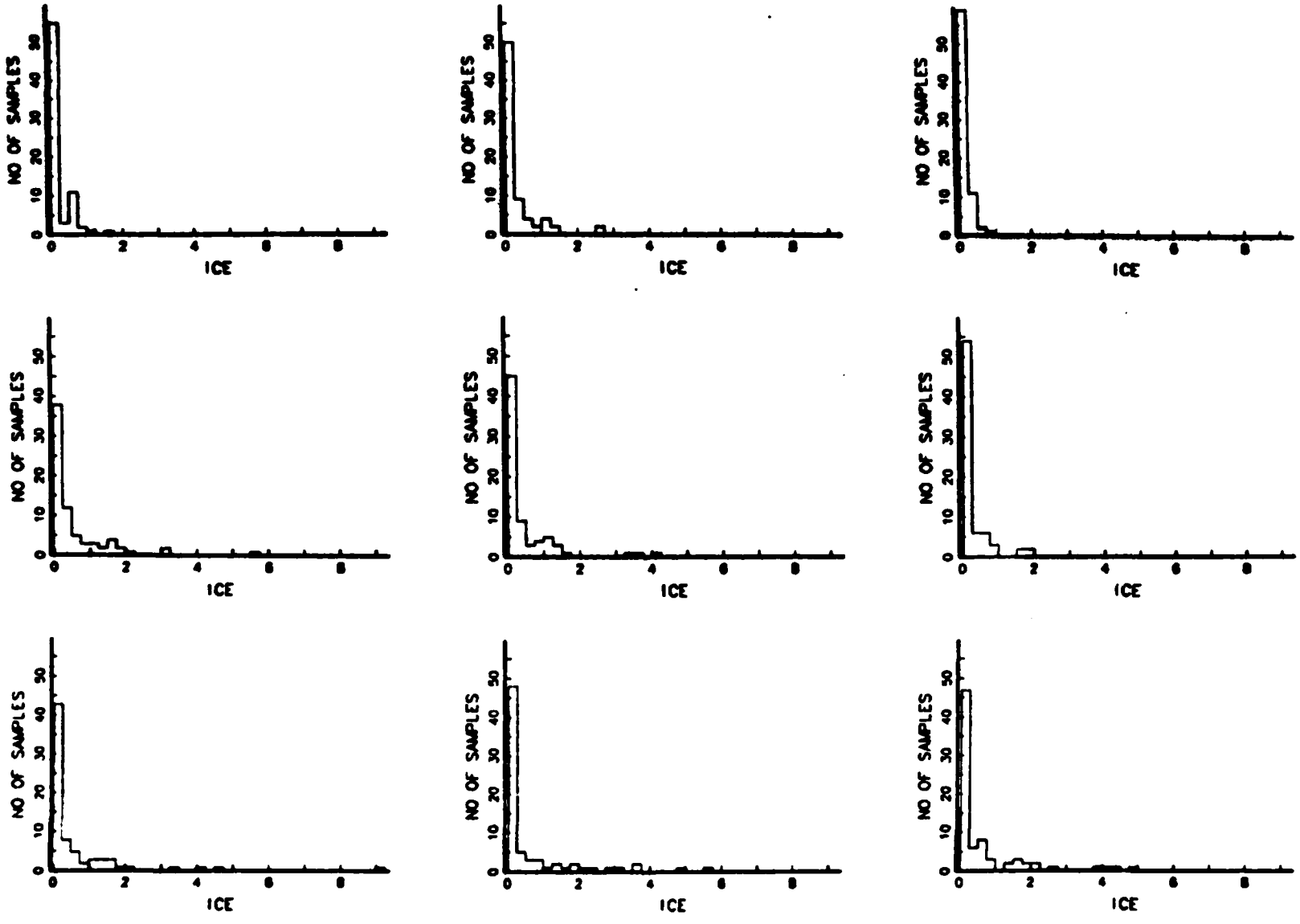


FIGURE 3

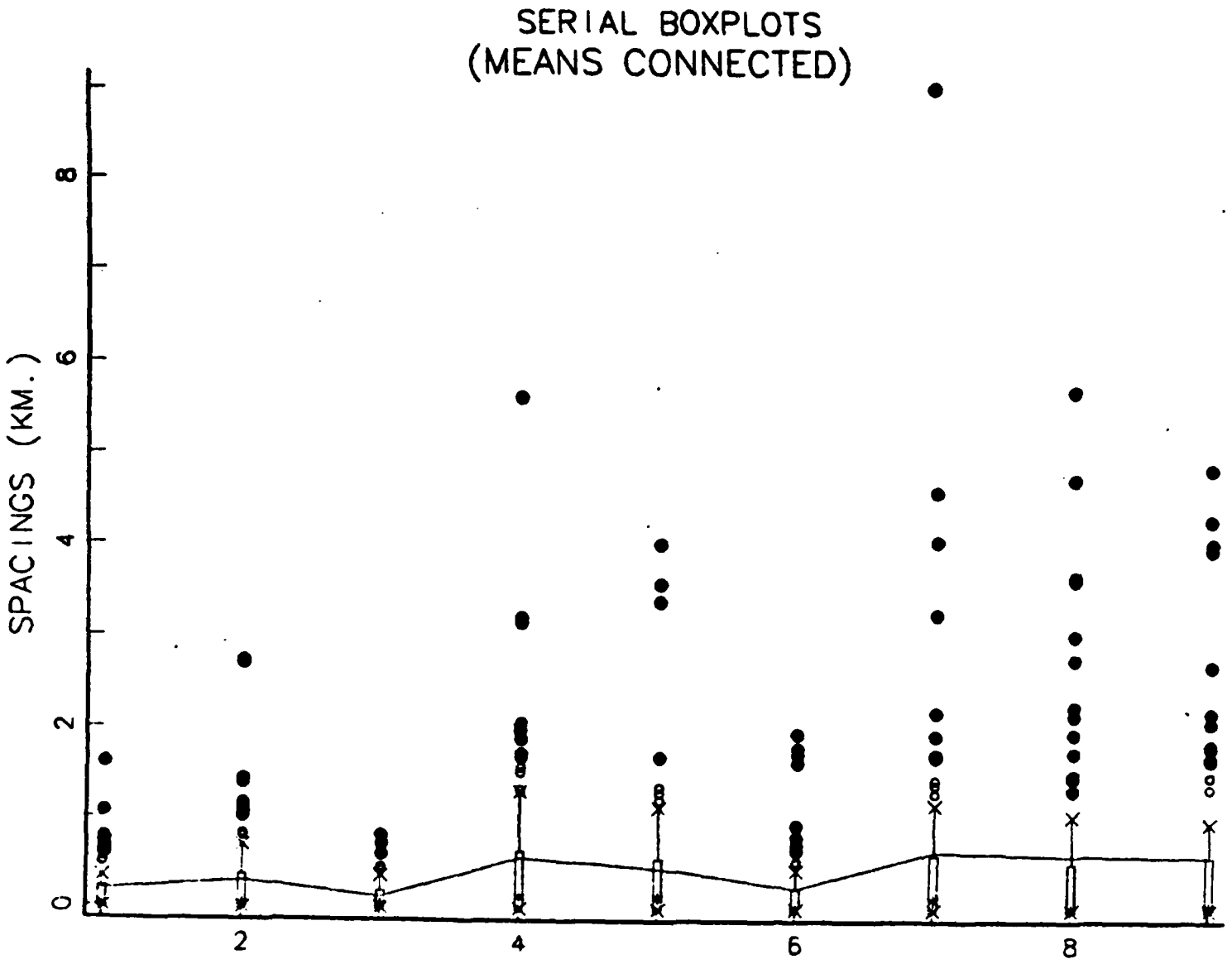


FIGURE 4

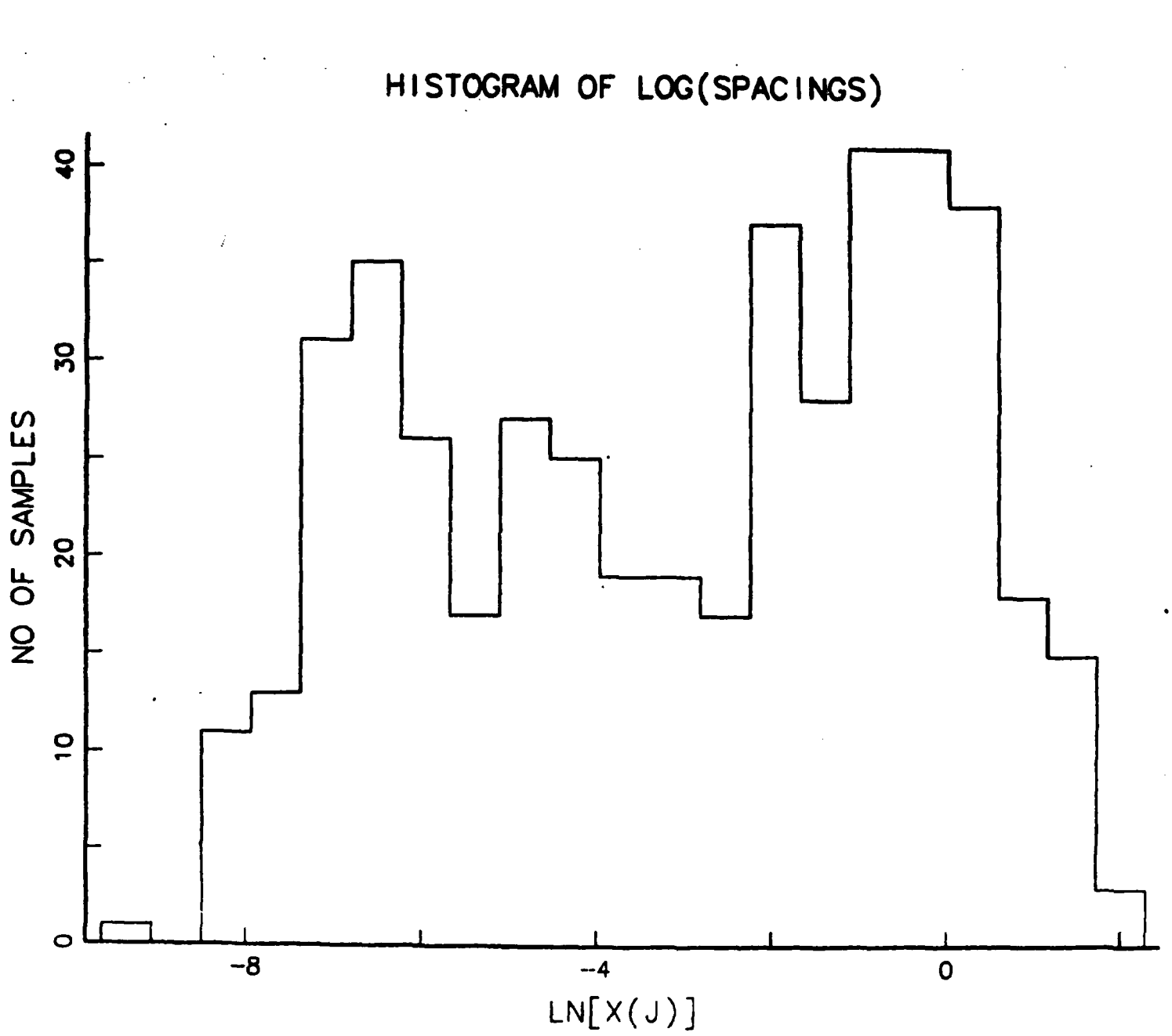


FIGURE 5

SERIAL HISTOGRAMS OF LOGGED SPACINGS

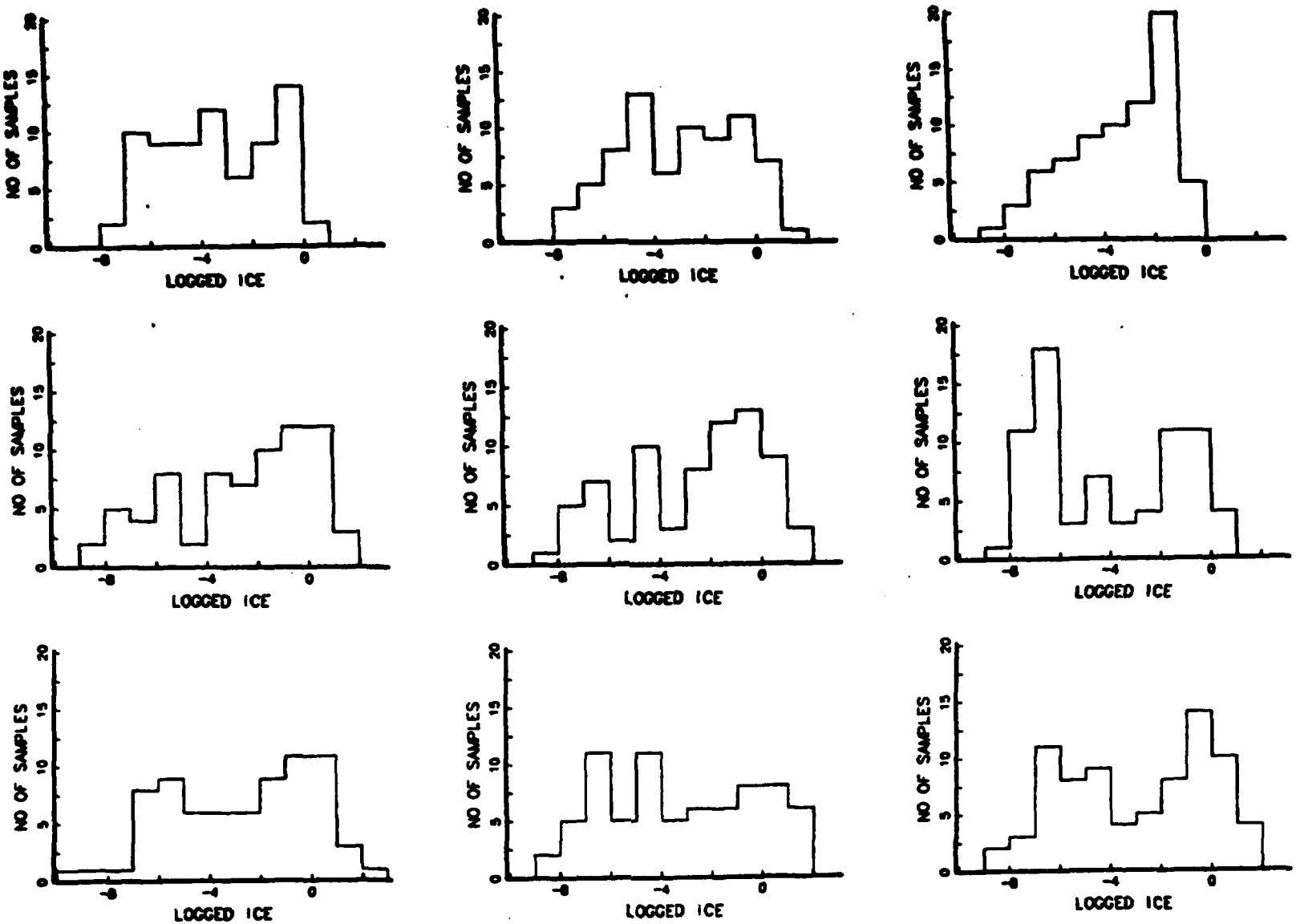


FIGURE 6

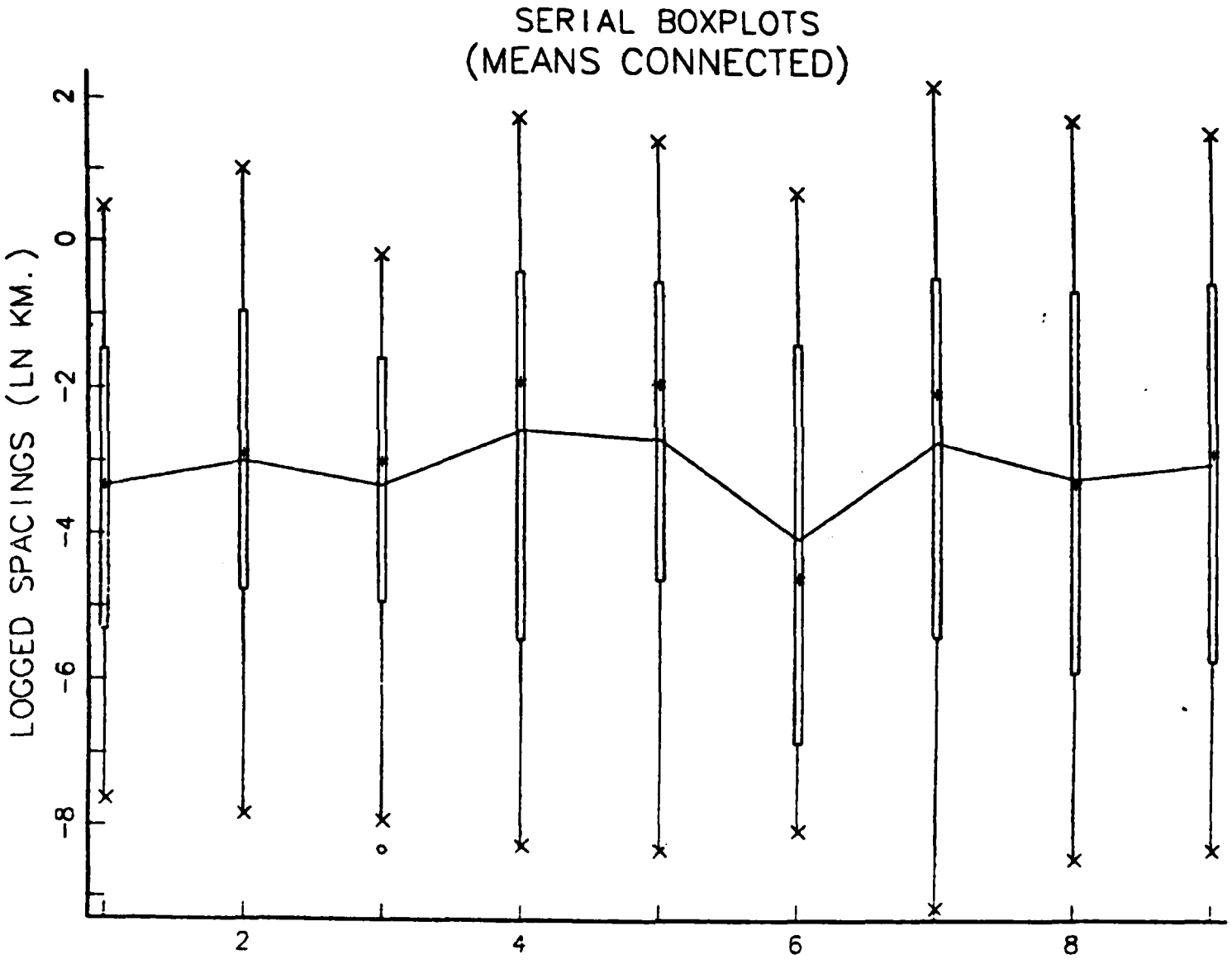

SERIAL BOXPLOTS
(MEANS CONNECTED)
LOGGED SPACINGS (LN KM.)
2
0
-2
-4
-6
-8
2
4
6
8

FIGURE 7

5. Towards a Parametric Description of Spacings

A simple exponential distribution summarization of spacings data has been suggested by Hibler [1972], and adopted by WH. also for purposes of discussion.

Model 1: $X \sim$ Exponential. Let X be a random variable representing a typical spacing. Then if x is any positive number, the simple exponential model is that

$$P\{X > x\} = e^{-\mu x} , \tag{5.1}$$

so the probability density of spacings is

$$f_X(x;\mu) = e^{-\mu x}\mu \qquad x \geq 0 \tag{5.2}$$

and μ is the rate of occurrence of spacings per unit distance; equivalently,

$$E[X] = \frac{1}{\mu} . \tag{5.3}$$

This says that the population average is $1/\mu$. It is well known that the maximum likelihood estimate of μ in model (5.1) is simply $\hat{\mu} = (\bar{x})^{-1}$, the inverse of the average of observations on X, supposing that successive spacings are identically distributed and independent. A time series (lagged correlation and spectrum) analysis of successive spacings at d = 30 ft. gives evidence of only very weak dependence between spacings; such dependence will be ignored in what follows.

An informal but informative check for the suitability of the exponential model is to examine a plot of the order statistics $x_{(j)}$ of the (upper half of the) data to the corresponding expected exponential order statistics: a straight line relationship signifies that the simple exponential fits

well. Examination of Figure 8 shows, and Figure 9 reinforces, the appearance of a systematic upward bow in the data, signifying a systematically longer-than-exponential right tail. Such an effect is also present for d = 40 ft., and d = 50 ft., but the curvature becomes progressively less noticeable as the depth, d , increases. It has been noted also that as depth decreases (eg. d = 10 ft.), there is a tendency for bowing to occur in the opposite direction, i.e. for spacings to become shorter-tailed than exponential.

A numerical summary of observed data characteristics, in addition to the message of Figures 8 and 9, thus suggests the need for a representation other than the exponential; alternatives are considered in the following sections.

Table 1.

Moment and Quantile Summaries of
Spacings in Excess of Median (70 meters) at Depth d = 30 ft.

Mean: $\bar{x} = 0.929$ (km.)

$s^2 \equiv \hat{var}[X] = 1.485$, $s = 1.218$

Coeff. of variation = $\frac{s}{\bar{x}} = 1.311$ (1.0)*

Skewness = $\hat{\gamma}_1 = 2.678$ (2)*

Kurtosis = $\hat{\gamma}_2 = 9.853$ (6)*

Lower Quartile $\equiv \underline{Q} = 0.155$

Median = 0.487

Upper Quartile $\equiv \overline{Q} = 1.279$

*Numbers in parentheses are characteristic of an Exponential model.

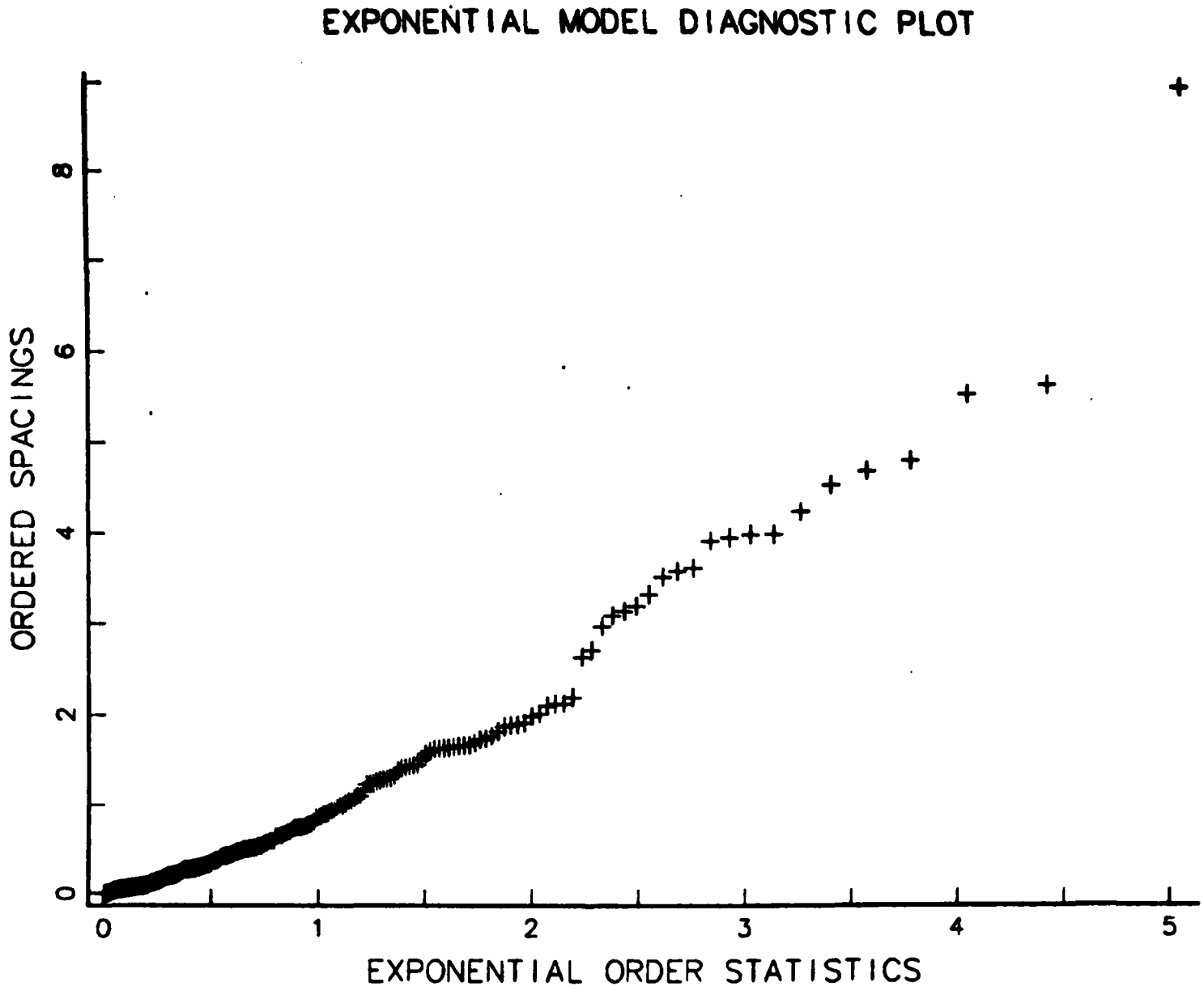

EXPONENTIAL MODEL DIAGNOSTIC PLOT
ORDERED SPACINGS
0
2
4
6
8
0
1
2
3
4
5
EXPONENTIAL ORDER STATISTICS

FIGURE 8

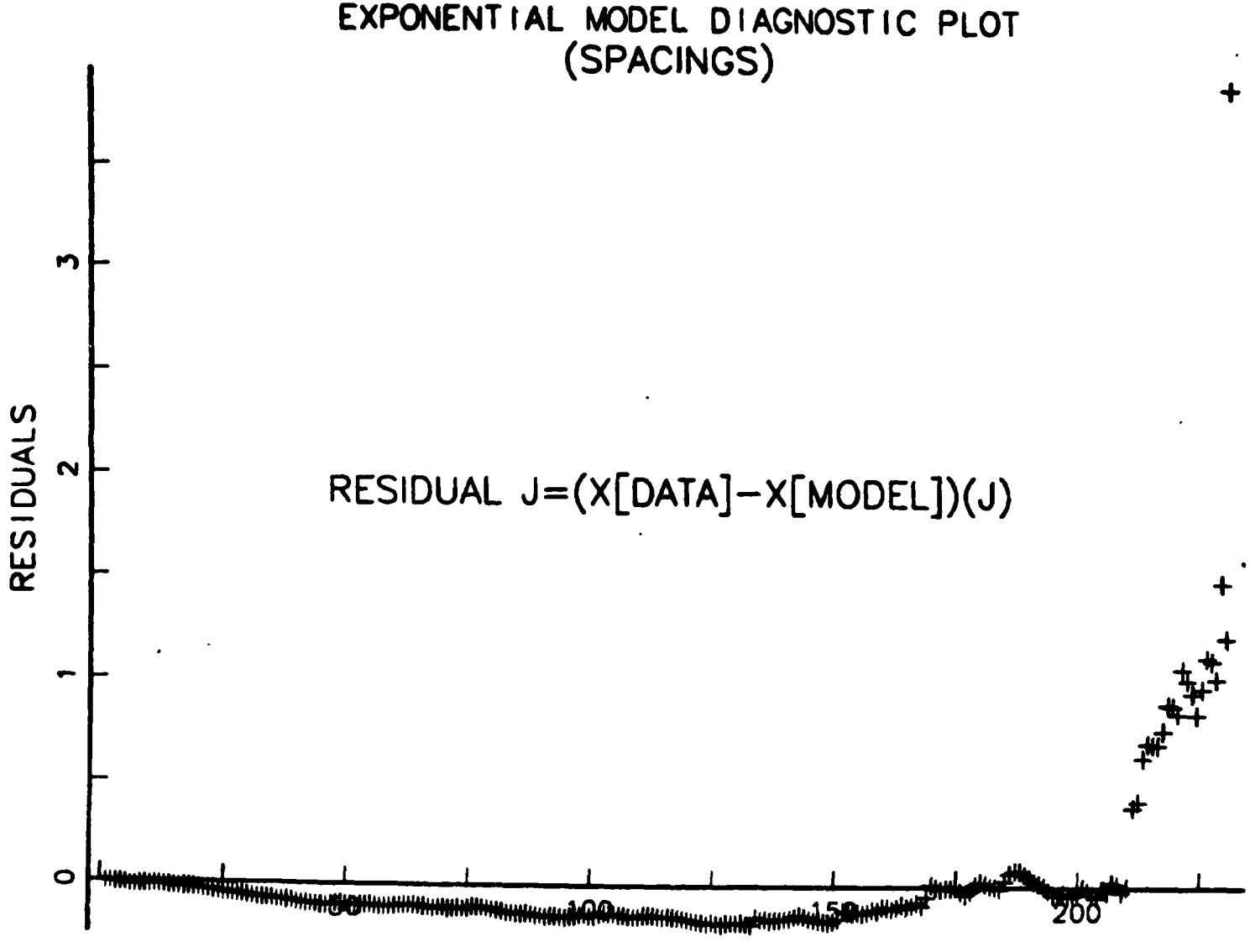

EXPONENTIAL MODEL DIAGNOSTIC PLOT
(SPACINGS)
RESIDUAL J=(X[DATA]-X[MODEL])(J)
RESIDUALS
0
1
2
3
100
150
200

FIGURE 9

The occurrence of a near-exponential distribution of spacings between deep keels is perhaps not surprising, in that ice structure formation appears to have a random nature without much long-term order. It is the near-independence of ice structure sizes in neighboring parts of the pack that is probably responsible for the near-exponentiality of the spacings observed. It is interesting that the spacing between two consecutive fixed high level upcrossings in certain Gaussian processes can be shown to be approximately exponential; see Cramér and Leadbetter [1967], Chap. 12. Of course ice depths are by no means Gaussian, but the conditions of weak long-run dependence relied upon to produce the Gaussian result are approximately present for the ice data as well. See also results of Gaver and Jacobs [1981] regarding the probability of reaching a high level in a non-Gaussian process. Agreement with the exponential distribution, both for spacings and keel depths, seems to improve with increase in reference depth, d, as is to be expected; numerical and graphical results are not, however, given in this report.

6. Modification of the Exponential Model: The Sculptured Exponential

We propose to fit the upper half of the d = 30 ft. spacings data, i.e. the magnitude of the spacing that exceeds the median spacing (about 70 m.) by a modified or sculptured exponential.

Model 2: Linearly Sculptured Exponential. As an alternative to the exponential (5.1) set

$$X = AZ(1 + CZ) , \tag{6.1}$$

where X represents a spacing, Z is a unit exponential basic r.v. and A and C are constants, A being a scale and C reflecting departure from exponentiality. The term $(1 + CZ)$ "sculptures" Z by leaving small values of Z virtually unchanged ($1 + CZ \simeq 1$ for Z small), but expanding large values ($1 + CZ \simeq CZ$ for Z large). Since X is represented as a monotonic increasing function of Z, we can represent the order statistics, $X_{(j)}$, as follows:

$$X_{(j)} = AZ_{(j)}(1 + CZ_{(j)}) , \tag{6.2}$$

that is the size-ordered X-values, $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ are easily represented in terms of those for Z, $Z_{(j)}$. Furthermore, if $Z_{(j)}$ is the jth order statistic of a unit exponential, then

$$Z_{(j)} = \frac{e_1}{n} + \frac{e_2}{n-1} + \ldots + \frac{e_j}{n-j+1} , \tag{6.3}$$

so the basic exponential $Z_{(j)}$ is represented as a weighted sum of exponential gaps; here $\{e_i\}$ is a sequence of iid exponential random variables. Even imbedding representation (6.3) into (6.2) or (6.1) is not difficult; exact

expectations of $X_{(j)}$'s are easily found. Note that, similarly, the quantiles or percent points, $x(\alpha)$, of X can be written in terms of those of Z, $z(\alpha)$:

$$x(\alpha) = Az(\alpha)(1 + Cz(\alpha)) , \qquad (0 \leq \alpha \leq 1) . \tag{6.4}$$

Thus sculpturing gives a simple representation of the inverse distribution of X in terms of that of Z.

The above considerations suggest that sculpturing is sometimes a natural way of fitting a Wilk-Gnanadesikan [1968] q-q plot. For further discussion see Gaver and Acar [1979], and Gaver [1982].

Expressions for the moments and (Pearsonian) skewness and kurtosis of (6.1) are obtainable from the following formulas:

$$E[X^k] = A^k \sum_{j=0}^{k} \binom{k}{j} (j+k)!C^j ; \qquad k = 1,2,3,... \tag{6.5}$$

in particular

$$E[X] = A(1+2C) , \; E[X^2] = A^2(2+2\times 3!C + 4!C^2) \tag{6.6}$$

$$E[X^3] = A^3[3! + 3\times 4!C + 3\times 5!C^2 + 6!C^3 ,$$
$$E[X^4] = A^4[4! + 4\times 5!C + 6\times 6!C^2 + 4\times 7!C^3 + 8!C^4] . \tag{6.7}$$

From these expressions there follow formulas for central moments obtained by substitution into the usual general formulas:

$$\text{Var}[X] = E[X^2] - (E[X])^2 \tag{6.8}$$

$$\text{Skew } X = \frac{E[X^3] - 3E[X^2]E[X] + 2(E[X])^3}{(\text{Var}[X])^{3/2}} \tag{6.9}$$

$$\text{Kurt}[X] = \frac{E[X^4] + 6E[X^2](E[X])^2 - 4E[X^3]E[X] - 3(E[X])^4}{(\text{Var}[X])^2} \,. \tag{6.10}$$

Furthermore, explicit formulas for the distribution and density functions of Model 2, (6.1), can be derived.

$$F_X(x;\, A,C) = P\{X \le x\} = 1 - \exp\left[-\frac{2x}{A + \sqrt{A^2 + 4ACx}}\right] \tag{6.11}$$

is the distribution function, and

$$f_X(x;\, A,C) = \exp\left[-\frac{2x}{A + \sqrt{A^2 + 4ACx}}\right] \cdot \frac{1}{\sqrt{A^2 + 4ACx}} \tag{6.12}$$

is the density.

The distributional form resembles that of an ordinary exponential with scale A and shape parameter 1 for small x, gradually transitioning into a longer-tailed Weibull with shape parameter 1/2 as x increases. A modification of the cumulative hazard in (6.11) suggests itself: instead of $2x/(A + \sqrt{A^2 + 4ACx})$ consider more generally $\lambda x/\{\alpha + (1-\alpha)(1+\gamma x^p)^q\}$; this leads to the distributional form

$$F_X(x;\lambda,\alpha,\gamma,p,q) = 1 - \exp\left[\lambda x/\{\alpha + (1-\alpha)(1+\gamma x^p)^q\}\right] \,, \tag{6.13}$$

$$p,q,\lambda,\gamma \ge 0 \,, \quad 0 \le \alpha \le 1 \,.$$

Note that (6.13) allows the representation of both ultimately long-tailed and ultimately short-tailed distributions relative to the exponential, but maintains exponential-like behavior for small x. Unfortunately, simple structural behavior of random variables akin to (6.1) disappears. The properties of distributions such as (6.13) will be the subject of future exploration; they are not considered further in this report.

Model 3: Exponentially Sculptured Exponential. A further alternative to the simple exponential (5.1) is of the form

$$X = AZe^{CZ}, \qquad A, C \geq 0 . \tag{6.14}$$

The sculpturing term e^{CZ} again leaves small values of Z (exponential) nearly unchanged, but considerably extends large Z values. Once again by monotonicity we have

$$X_{(j)} = AZ_{(j)}e^{CZ_{(j)}} \qquad (j = 1,2,\ldots, n) \tag{6.15}$$

and

$$x(\alpha) = Az(\alpha)e^{Cz(\alpha)}, \qquad (0 \leq \alpha \leq 1) \tag{6.16}$$

for order statistics and quantiles of X from (6.14).

Expressions for the moments of this model come by differentiating the moment-generating function of Z,

$$E[e^{\theta Z}] = (1-\theta)^{-1}, \qquad (0 \leq \theta < 1);$$

we obtain

$$E[X] = A(1-C)^{-2} \qquad (C < 1) \qquad (6.17)$$

$$E[X^2] = A^2(1-2C)^{-3} \qquad (C < 1/2) \qquad (6.18)$$

$$E[X^3] = A^3(1-3C)^{-4} \qquad (C < 1/3) \qquad (6.19)$$

$$E[X^4] = A^4(1-4C)^{-5} \qquad (C < 1/4) \qquad (6.20)$$

The central moments, i.e. variance, skewness, and kurtosis, then come with the aid of (6.8), (6.9), and (6.10).

It is not possible to derive a simple expression for the distribution function and density of the model (6.14).

7. Fitting the Sculptured Exponential Model for Spacings

A number of ways of fitting models or representations such as (6.1) and (6.14) suggest themselves; among these are the following.

(A) Maximum likelihood: possible for (6.1), using (6.12).

(B) Moment-matching: possible for (6.1), and also (6.14).

(C) Quantile matching: feasible for any sculptured representation.

(D) Hybrid methods: e.g. constrained likelihood fit by requiring $E[X] = \bar{x}$ and allowing C in (6.1) to be determined by maximum likelihood.

(E) Generalized non-linear least-squares, robustified if necessary: it is proposed to regress $x_{(j)}$ on $Az(j)[1 + Cz(j)]$ where e.g. $z(j) = -\ell n(1-j/(n+1))$, the approximate expected value of the basic r.v. Z, and an appropriate covariance matrix is utilized.

We report the results of applying several of these methods to fit Models 2 and 3 to spacings (at d = 30 ft.) data; the results are then diagnostically examined.

(A) Maximum Likelihood.

The log likelihood function associated with (6.1) and (6.12) is

$$L(A,C;\ \underline{x}) = -\sum_{j=1}^{n}\left\{\frac{2x_j}{A + \sqrt{A^2 + 4AC\ x_j}} - \frac{1}{2}\,\ell n(A^2 + 4AC\ x_j)\right\}. \qquad (7.1)$$

Differentiation leads to likelihood equations for $\hat{A}$ and $\hat{C}$; these must be solved numerically, e.g. by a Newton-Raphson method. Alternatively, a search of the likelihood function itself is reasonably effective. Results are summarized below for the 30 ft. depth data set.

Table 2

Maximum Likelihood Fits of Spacings by a

Sculptured Exponential, (6.1).

$n = 231$, $\hat{A} = 0.53(0.09)^*$, $\hat{C} = 0.39(0.16)^*$

$(\text{Corr}(\hat{A},\hat{C}) = -0.84)^*$

Estimates	Raw Data	Model
$\hat{E}[X]$	0.93	0.95
Coeff. $\hat{V}$ar.[X]	1.31	1.50
$\hat{S}$kew[X]	2.68	4.57
$\hat{K}$urt[X]	9.85	40.32
Lower Quartile, $\underline{Q}$	0.155	0.170
Median	0.487	0.467
Upper Quartile, $\overline{Q}$	1.28	1.13

*( ) represent large-sample standard errors calculated from the likelihood (Fisher information).

Although the agreement of the lower moments is satisfactory, that of the higher moments (skewness and kurtosis) is much less so; this points to the apparent sensitivity of the m.l.e. to extreme values, assigning an unreasonably high C ("correction") value. Nevertheless, a Kolmogorov-Smirnov test of goodness of fit yields a value of 0.65 for the fitted sculptured model, while a Kolmogorov-Smirnov value of 1.73 is found for a simple exponential fit; the sculptured model is seen by this test as providing a substantially improved fit. Furthermore, a diagnostic plot of observed (x_j) vs. predicted $(\hat{x}_j(\hat{A},\hat{C})$ order statistics provides a more satisfactory straight-line fit than does an exponential model. See Figure 10. A further plot of residuals is given below in Figure 11. The sculptured model fitted by maximum likelihood to the particular set of data under discussion appears to predict a somewhat longer far right tail than is evident from the data.

An alternative diagnostic plot is available for the present and certain other sculptured models: for (6.1) the transformation obtained by solving $x = Az(1 + Cz)$ for z, using the estimated parameter values provides estimates of the basic z-values, denoted by $\hat{z}$, giving rise to the observations. To the extent that the latter resemble observations on a unit exponential, the model is verified. Figures 12 and 13 provide such a diagnostic plot.

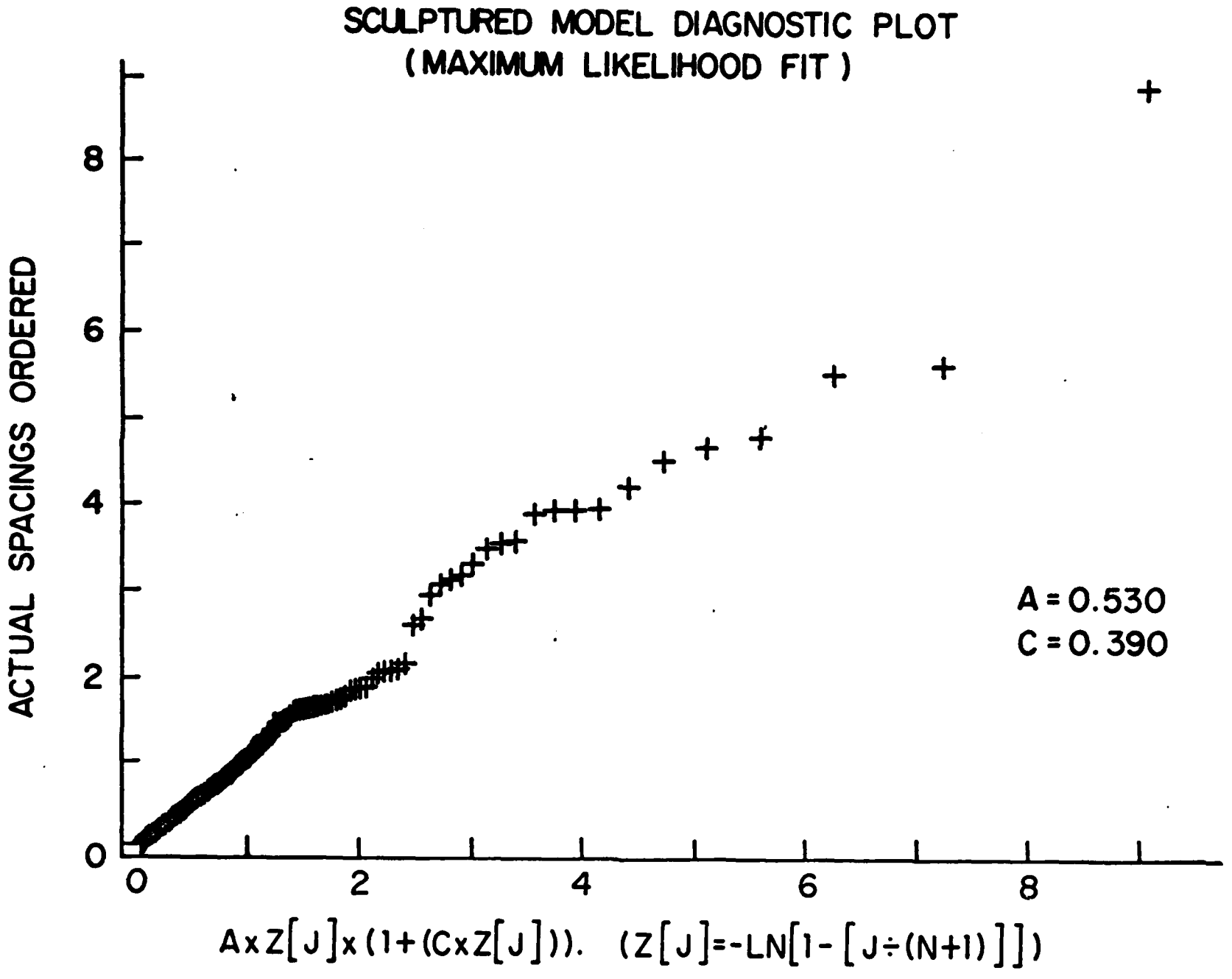

SCULPTURED MODEL DIAGNOSTIC PLOT
(MAXIMUM LIKELIHOOD FIT)
ACTUAL SPACINGS ORDERED
0
2
4
6
8
A = 0.530
C = 0.390
A x Z[J] x (1+(C x Z[J])). (Z[J]=-LN[1-[J÷(N+1)]])

Figure 10

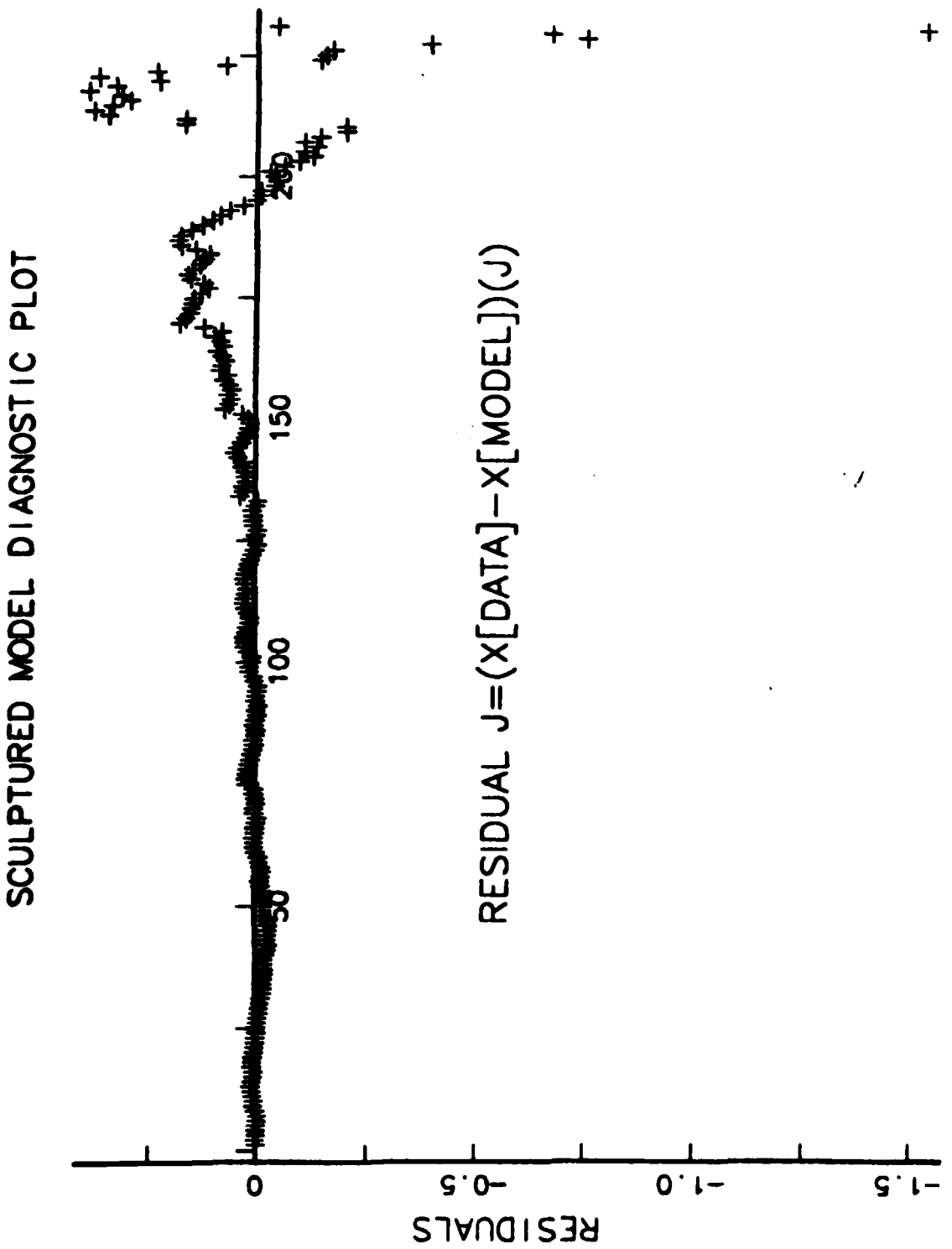

SCULPTURED MODEL DIAGNOSTIC PLOT
RESIDUAL J=(X[DATA]-X[MODEL])(J)
RESIDUALS
0
-0.5
-1.0
-1.5
50
100
150
200

FIGURE 11

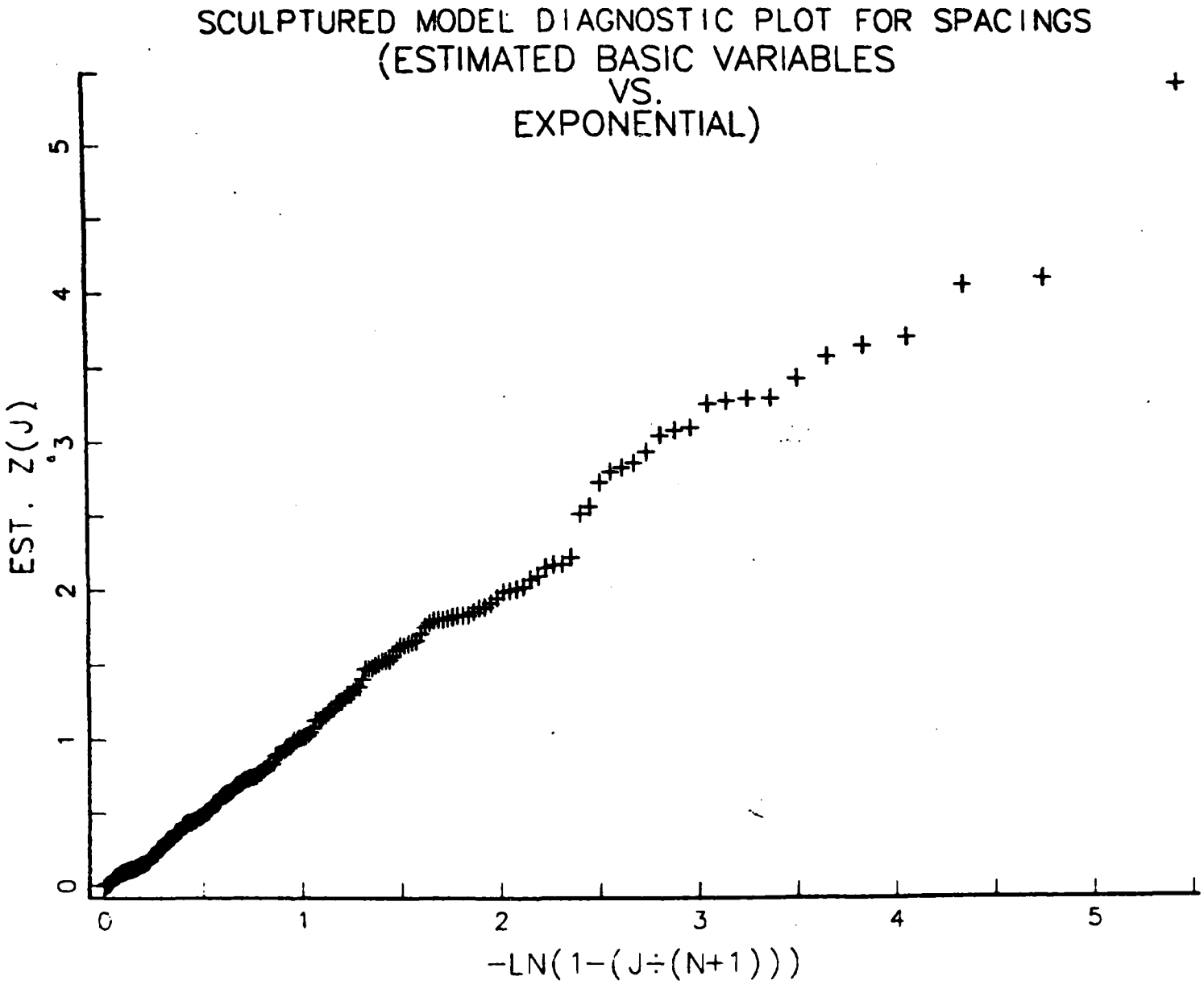

SCULPTURED MODEL DIAGNOSTIC PLOT FOR SPACINGS
(ESTIMATED BASIC VARIABLES
VS.
EXPONENTIAL)
EST. Z(J)
0
1
2
3
4
5
-LN(1-(J÷(N+1)))

FIGURE 12

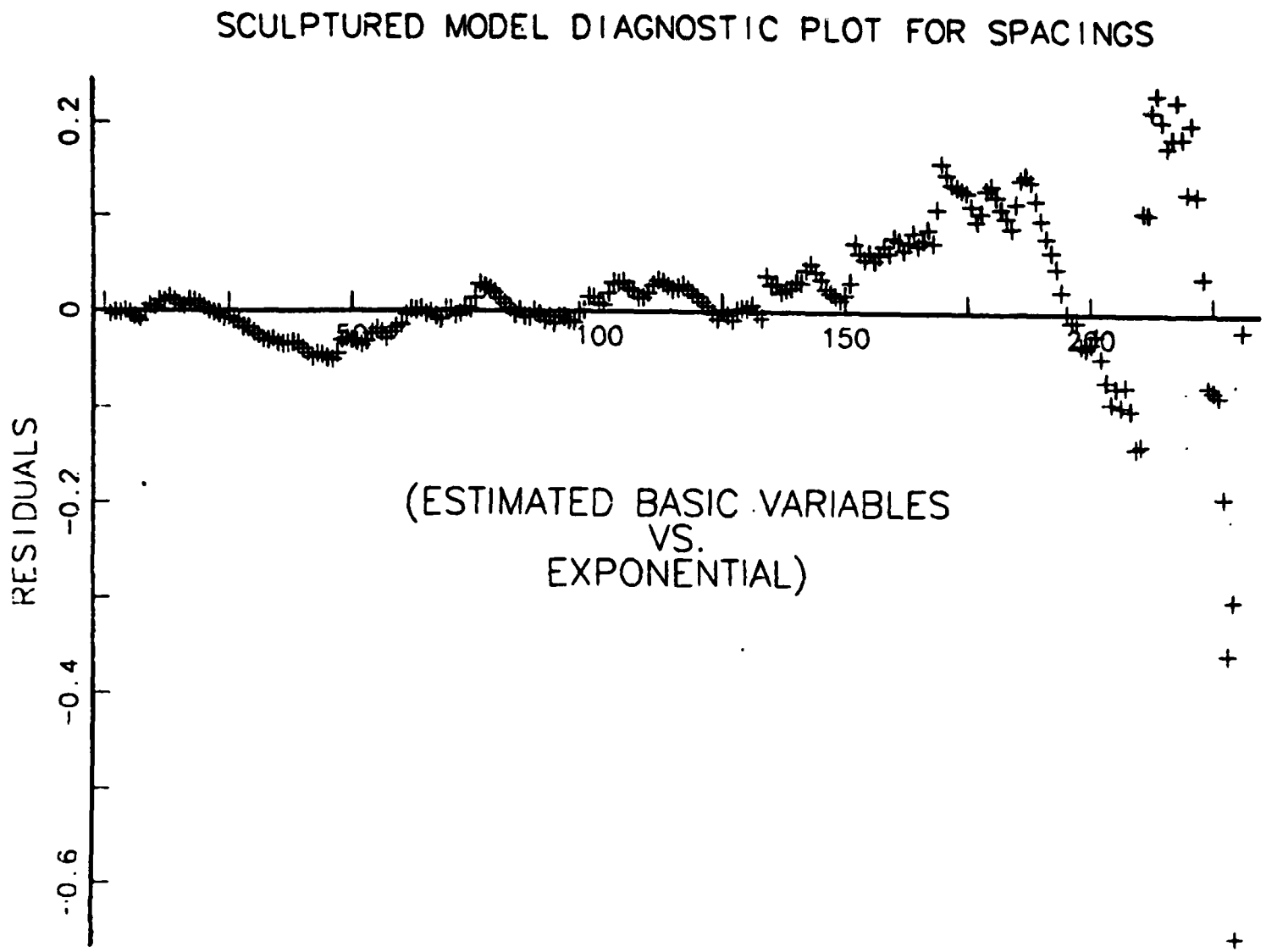

SCULPTURED MODEL DIAGNOSTIC PLOT FOR SPACINGS
RESIDUALS
0.2
0
-0.2
-0.4
-0.6
50
100
150
200
(ESTIMATED BASIC VARIABLES
VS.
EXPONENTIAL)

FIGURE 13

(B) Moment Matching: Linear Sculpturing

An alternative to the maximum likelihood method is that of matching moments. It is customary to equate the two lowest moments, e.g. sample and model mean and variance, when fitting two-parameter models. There is theoretical justification in the present case for matching sample and model means and coefficients of variations: the goodness-of-exponential-fit test of Stephens [1978], adapted from ideas of Shapiro and Wilk [1972], is essentially based on the sample of coefficient of variation; (cf. Shapiro and Wilk [1972], p. 357 footnote). The coefficient of variation is a simple rational function for the model (6.1), and $\tilde{C}$ can be found explicitly.

The skewness of model (6.1) is also a simple monotonic expression in C; this can be equated to the sample skewness and solved for an estimate of C , $\tilde{C}$. An estimate of A then is obtainable from the first moments.

When applied to spacings at d = 30 ft. the skewness-matching method produces estimates that differ noticeably from those given by maximum likelihood: $\tilde{A}$ is larger, and $\tilde{C}$ is smaller than the corresponding maximum likelihood estimates. In several respects the skewness-match is to be preferred: it agrees best with the data evidence in the far tail, i.e. at the upper quartile and with respect to the kurtosis measures of model and data.

Table 3

Moment - Matched Fits of Spacings

by A Sculptured Exponential

Estimates	Raw Data	Model C.V. Match $(\tilde{A} = 0.67, \tilde{\tau} = 0.193)$	Model Skew. Match $(\tilde{A} = 0.827, \tilde{\tau} = 0.062)$
$E[\hat{X}]$	0.929	0.929	0.929
$\hat{\text{Coeff. Var.}}[X]$	1.31	1.31	1.12
$\hat{\text{Skew}}[X]$	2.68	3.71	2.68
$\hat{\text{Kurt}}[X]$	9.85	26.01	12.54
Lower Quartile, $\underline{Q}$	0.155	0.203	0.242
Median	0.487	0.53	0.598
Upper Quartile, $\overline{Q}$	1.29	1.18	1.24

Figures 14 and 15 provide diagnostic plots of quality of fit for the skewness-matched sculptured model, (6.1).

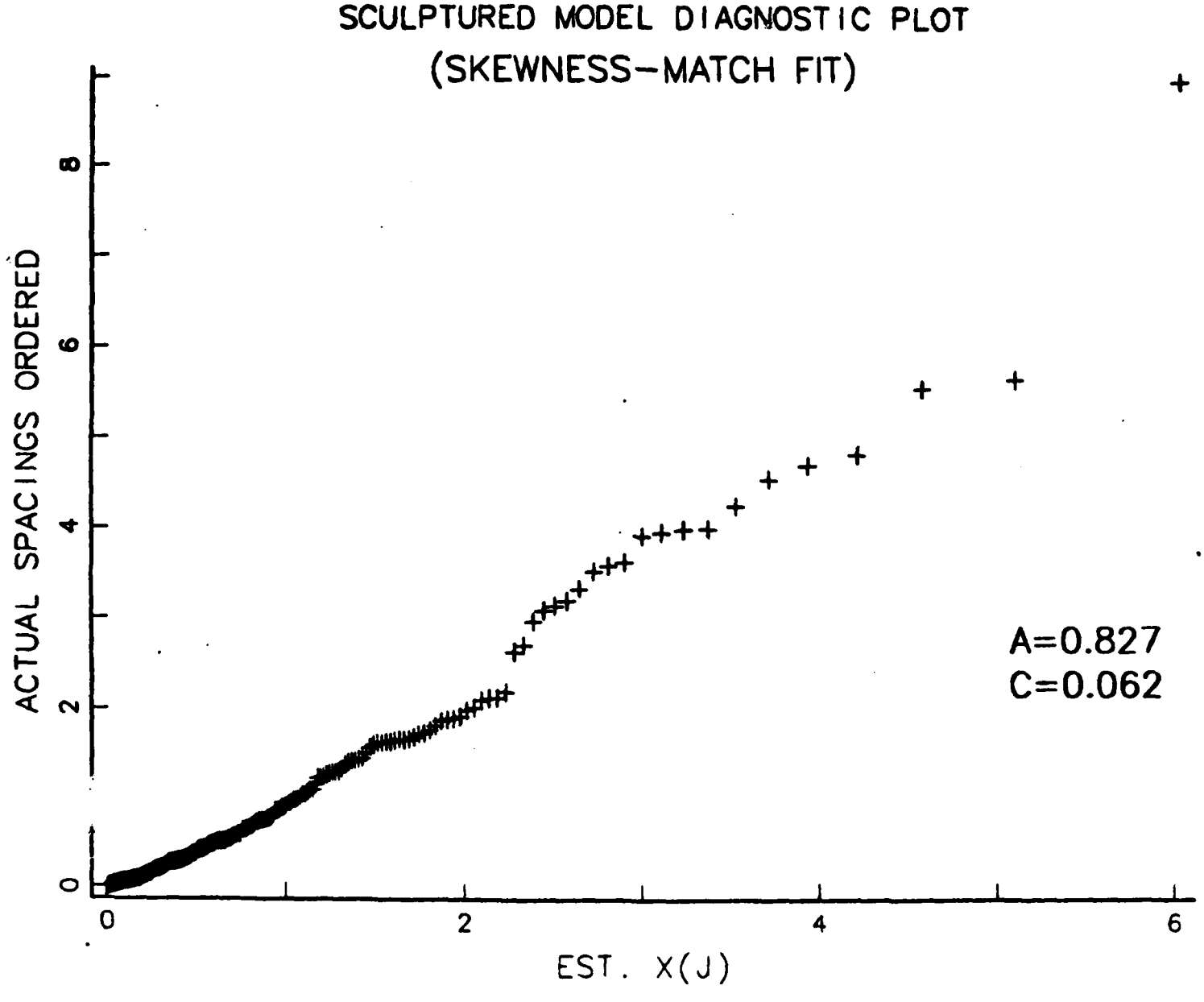

SCULPTURED MODEL DIAGNOSTIC PLOT
(SKEWNESS–MATCH FIT)
ACTUAL SPACINGS ORDERED
EST. X(J)
0
2
4
6
8
A=0.827
C=0.062

FIGURE 14

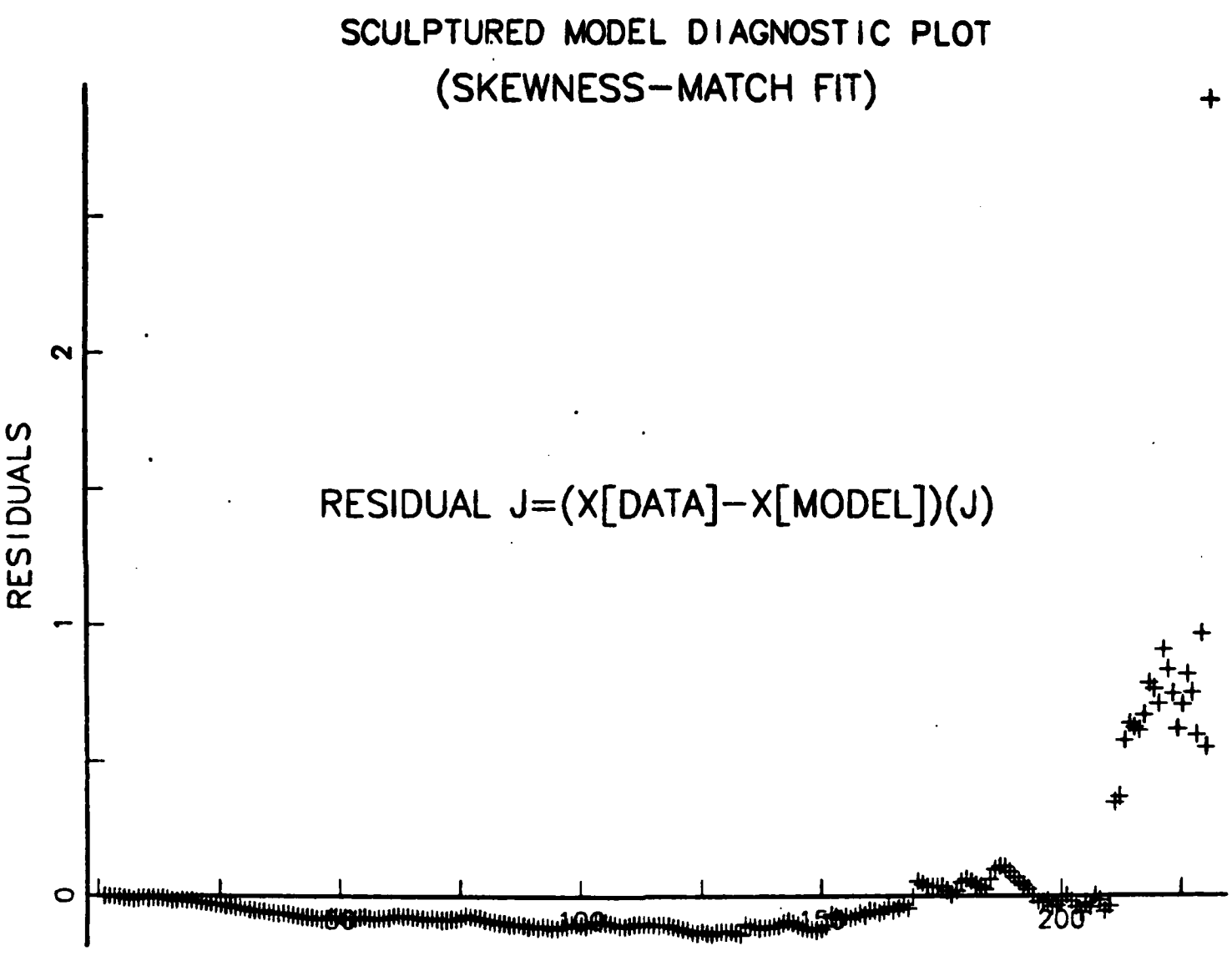

SCULPTURED MODEL DIAGNOSTIC PLOT
(SKEWNESS-MATCH FIT)
RESIDUAL J=(X[DATA]-X[MODEL])(J)
RESIDUALS
0
1
2
100
150
200

FIGURE 15

(C) Moment Matching: Exponential Sculpturing

The model (6.15) was also fit to the data by choosing C to match the coefficient of variation and choosing A to match the mean of the data. The values of the estimated parameters and predicted moments are as follows.

Table 4.

Coefficient of Variation - Match fit of Spacings by the Sculptured Model (6.14)

$\tilde{A} = 0.721$ $\tilde{C} = 0.119$

Estimates	Raw Data	Model
$\hat{E}[X]$	0.929	0.929
$\hat{CV}.[X]$	1.31	1.31
$\hat{Skew}[X]$	2.68	4.53
$\hat{Kurt}.[X]$	9.84	53.64
Lower Quartile, $\underline{Q}$	0.155	0.215
Median	0.487	0.543
Upper Quartile, $\overline{Q}$	1.29	1.18

Figures 16 and 17 provide diagnostic plots of quality of fit for the model (6.14). The residuals appear to have the same shape as the residuals for the m.l.e. and skewness-match fit of model (6.1). The residuals in Figure 17 are more symmetrically and closely grouped about the axis than are the residuals for the maximum likelihood fitted linear model.

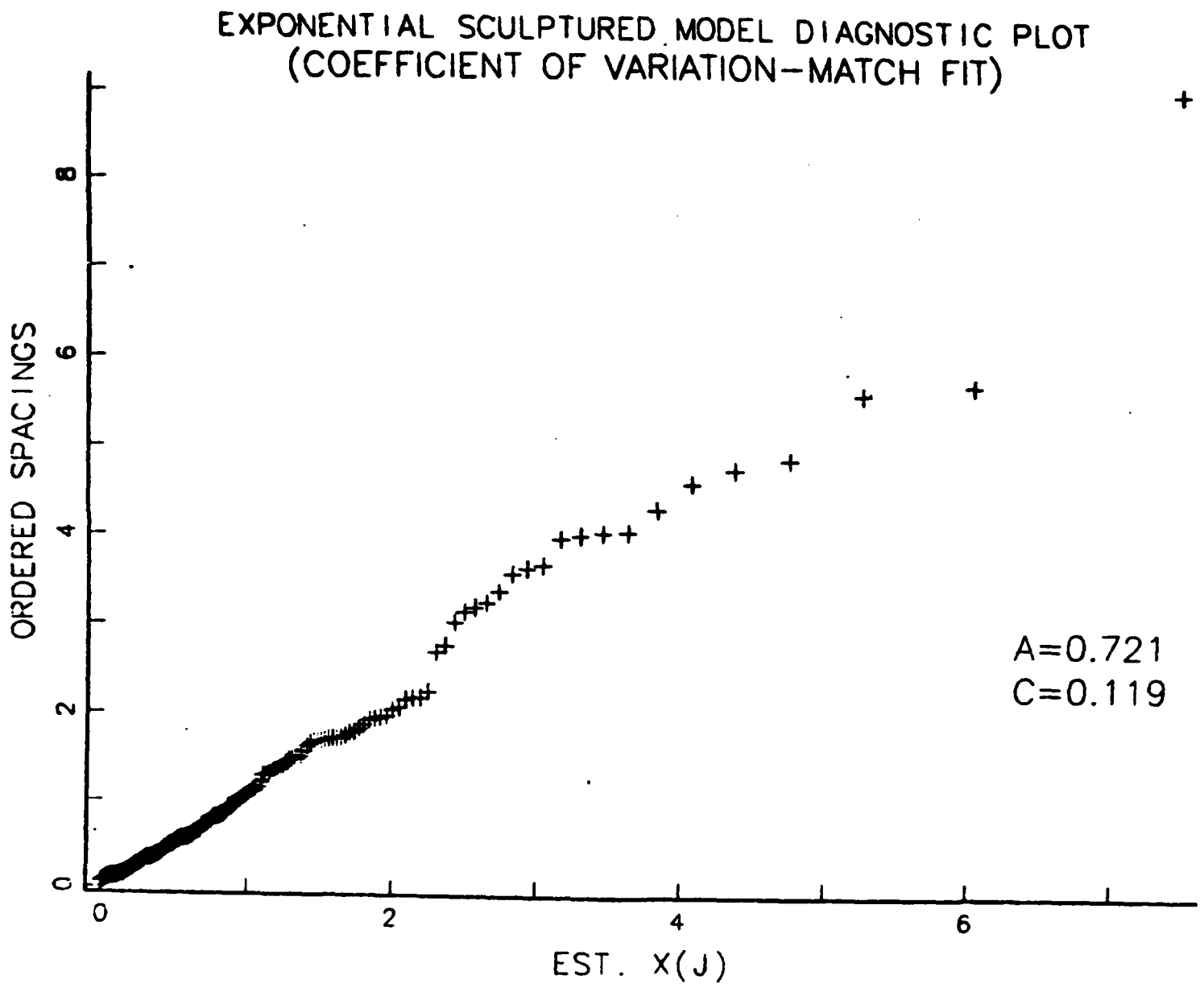

EXPONENTIAL SCULPTURED MODEL DIAGNOSTIC PLOT
(COEFFICIENT OF VARIATION–MATCH FIT)
ORDERED SPACINGS
0
2
4
6
8
EST. X(J)
A=0.721
C=0.119

FIGURE 16

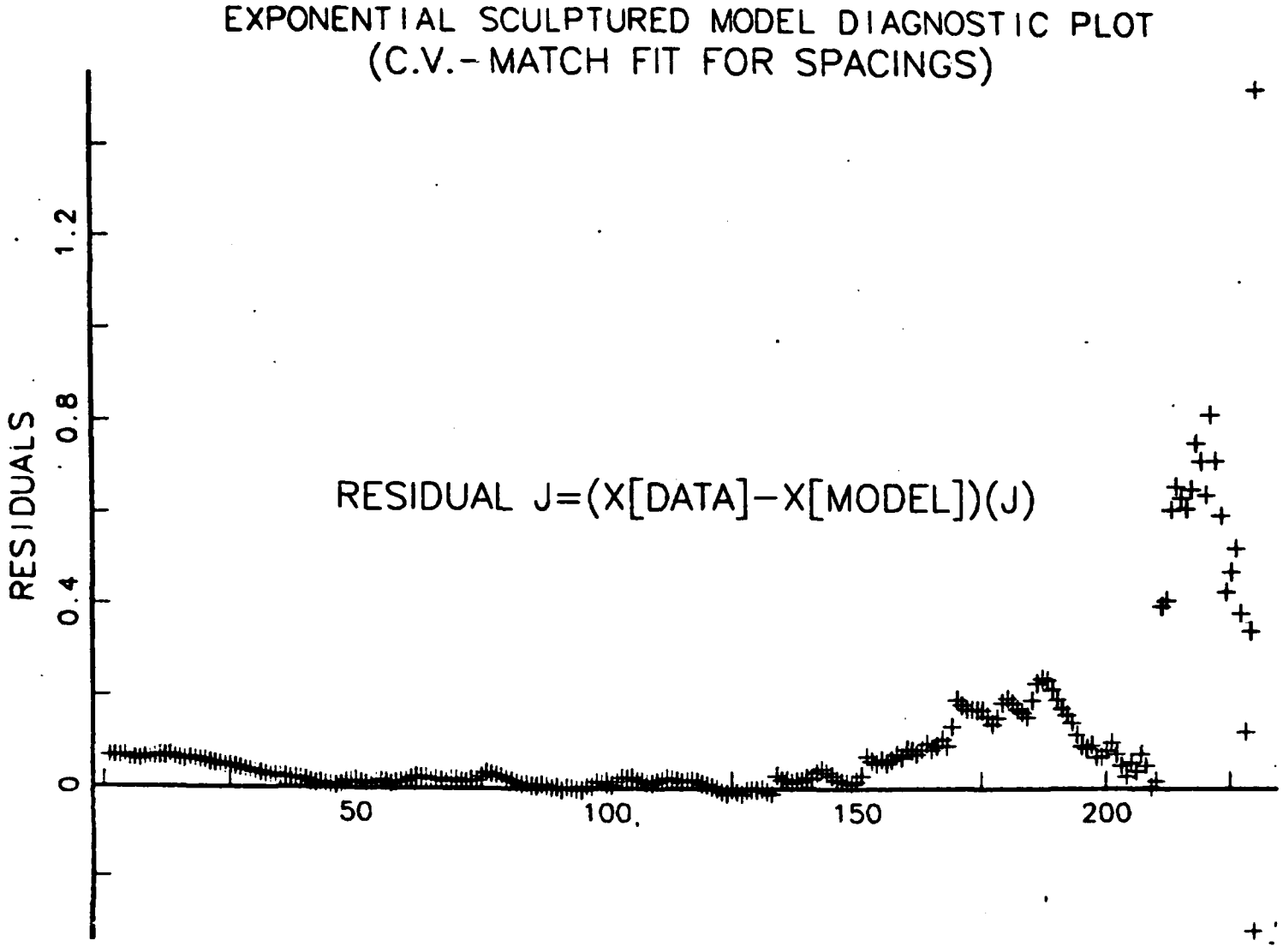

EXPONENTIAL SCULPTURED MODEL DIAGNOSTIC PLOT
(C.V.- MATCH FIT FOR SPACINGS)
RESIDUAL J=(X[DATA]-X[MODEL])(J)
RESIDUALS
0
0.4
0.8
1.2
50
100
150
200

FIGURE 17

8. The Gamma Model for Spacings

The gamma distribution is an alternative, and classical, model for spacings that includes the exponential as a special case. Its density is

$$f_X(x;\, B,A) = e^{-x/B}\,\frac{(x/B)^{A-1}}{\Gamma(A)}\cdot\frac{1}{B}\,. \tag{8.1}$$

Neither the distribution function nor the quantiles or order statistics of general gamma-distributed random variables are explicitly expressible in simple closed form.

The gamma has been fitted to the 30 ft. spacings data by maximum likelihood and also by matching the first two moments. The results are summarized below.

Table 5.

Estimates	Raw Data	Model(Moments) ($\tilde{A}$ = 0.582, $\tilde{C}$ = 1.60)	Model(M.L.E.) ($\tilde{A}$ = 0.70, $\tilde{B}$ = 1.33) S.E. (0.15) (0.60) Corr($\tilde{A}$,$\tilde{B}$) = - .71
$\hat{E}[X]$	0.93	0.929	0.929
$\hat{C}oeff.\ Var[X]$	1.31	1.311	1.196
$\hat{Skew}[X]$	2.68	2.622	2.390
$\hat{Kurt}[X]$	9.85	10.32	8.58
Lower Quartile, $\underline{Q}$	0.155	0.127	0.172
Median	0.487	0.478	0.541
Upper Quartile, $\overline{Q}$	1.279	1.256	1.276

In Figures 18 and 19 diagnostic plots of the gamma fit by maximum likelihood are presented. There is a remarkable resemblance between the appearance of the diagnostics for the m.l.e.- fitted gamma model and the skewness-fitted sculptured exponential for this data set. Note that both of these model representations now tend to predict smaller extreme right tails than indicated by the raw data, in contrast to the m.l.e.-fitted sculptured exponential.

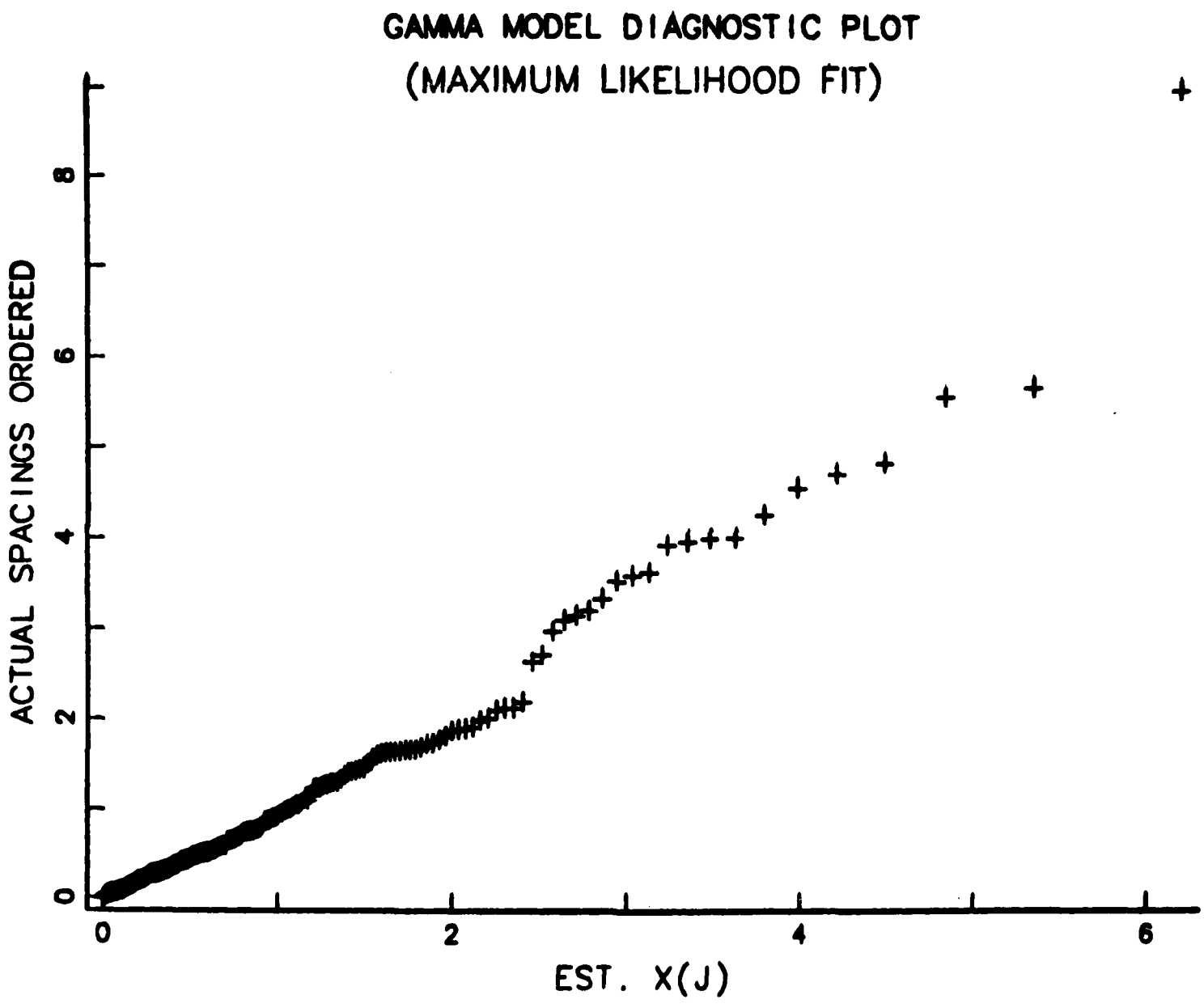

GAMMA MODEL DIAGNOSTIC PLOT
(MAXIMUM LIKELIHOOD FIT)
ACTUAL SPACINGS ORDERED
0
2
4
6
8
0
2
4
6
EST. X(J)

FIGURE 18

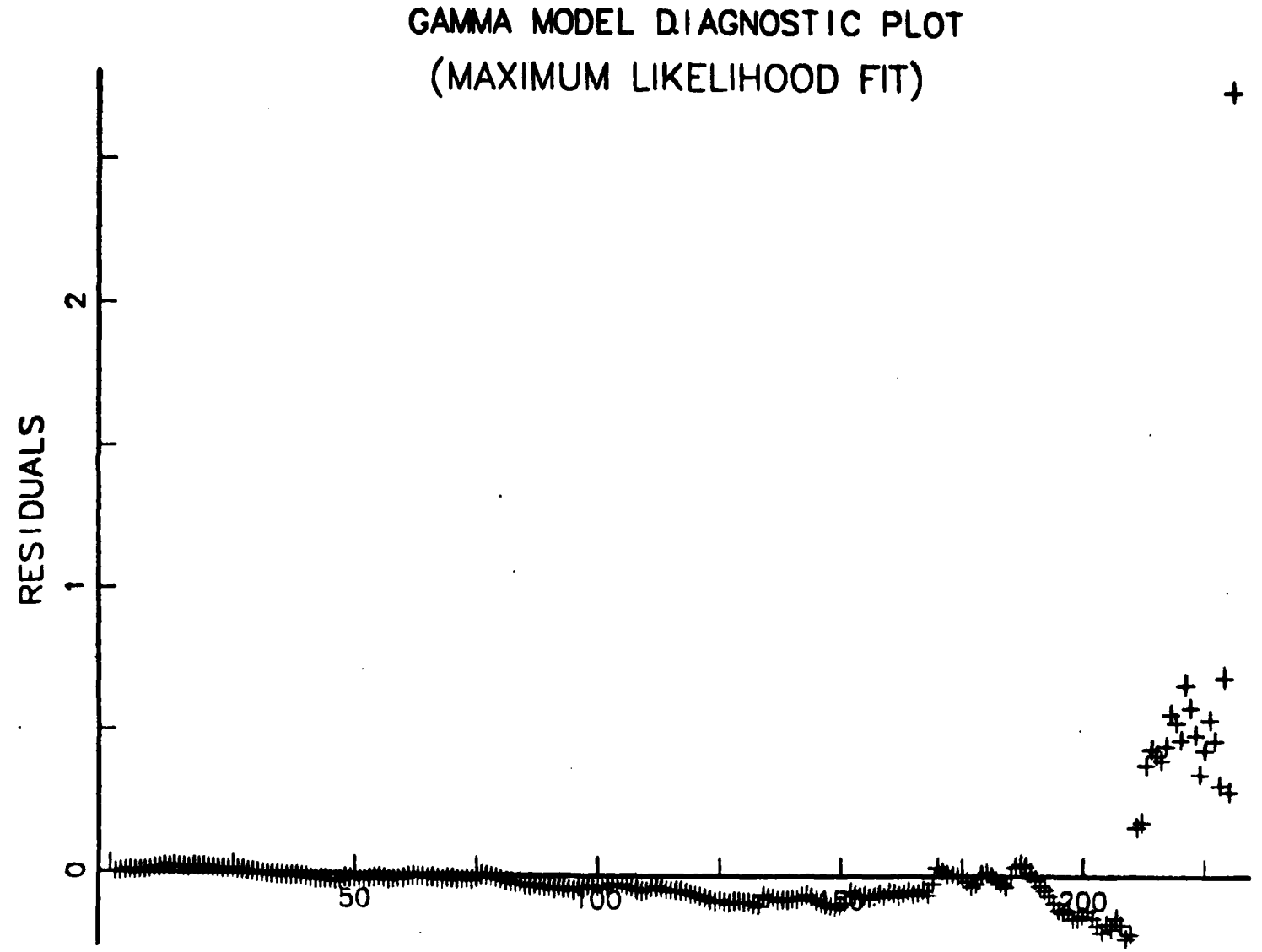

GAMMA MODEL DIAGNOSTIC PLOT
(MAXIMUM LIKELIHOOD FIT)
RESIDUALS
0
1
2
50
100
150
200

FIGURE 19

9. Statistical Properties of Ice Keels

Turn now to a discussion of the distributional properties of keels, i.e. the locally maximal projections of individual ice structures below the sea surface; see Figure 1. The present discussion is confined to those keels referenced from a 30 ft. depth (keel depths are in units of feet).

Figure 20 is a plot of keel depths (excess over 30 ft.) in sequential order, Figure 21 is a histogram of raw keel depths, and Figure 22 is a histogram of log keel depths. The latter shows little evidence of the pronounced "two-bump effect" of the spacings. Earlier work, WH, and Hibler [1972], among others, has represented keel depths by the simple exponential model, but again the data give evidence of a systematically longer-than-exponential right tail; see Figures 23 and 24.

Two theoretical models were fitted to the raw data: the sculptured exponential, (6.1), and the gamma. Several methods of fitting were employed: lower moment-matching and maximum likelihood. The adequancy of the fits was assessed by numerical and graphical methods, and the following tables and figures 25-30 summarize the results.

Table 6

Fits of Models to Ice Keel Depths Beyond 30 ft.
(n = 365)

Estimates	Raw Data	Sculptured Exponential			Gamma	
		M.L.E.	C.V.	Skew	M.L.E.	Moment
		$\hat{A}$=4.04 $\hat{C}$=0.37	$\tilde{A}$=5.27 $\tilde{C}$=.155	$\tilde{A}$=6.65 $\tilde{C}$=.019	$\hat{A}$=.73 $\hat{B}$=9.43	$\tilde{A}$=.63 $\tilde{B}$=11
		S.E. (.52) (.12)			S.E. (.83) (.04)	
		Corr($\hat{A}$,$\hat{C}$)=-.83			Corr($\hat{A}$,$\hat{B}$)=-.72	
$\hat{E}[X]$	6.90	7.03	6.90	6.90	6.90	6.90
$\hat{C.V.}[X]$	1.26	1.49	1.26	1.04	1.17	1.26
Skew[X]	2.22	4.51	3.47	2.22	2.34	2.53
Kurt[X]	5.55	39.32	22.4	7.90	8.20	9.57
$\underline{Q}$	1.20	1.29	1.58	1.92	1.36	1.08
Med.	3.60	3.52	4.04	4.67	4.11	3.77
$\overline{Q}$	8.60	8.47	8.88	9.46	9.48	9.44

K.S. for raw data to be exponential = 2.3 .

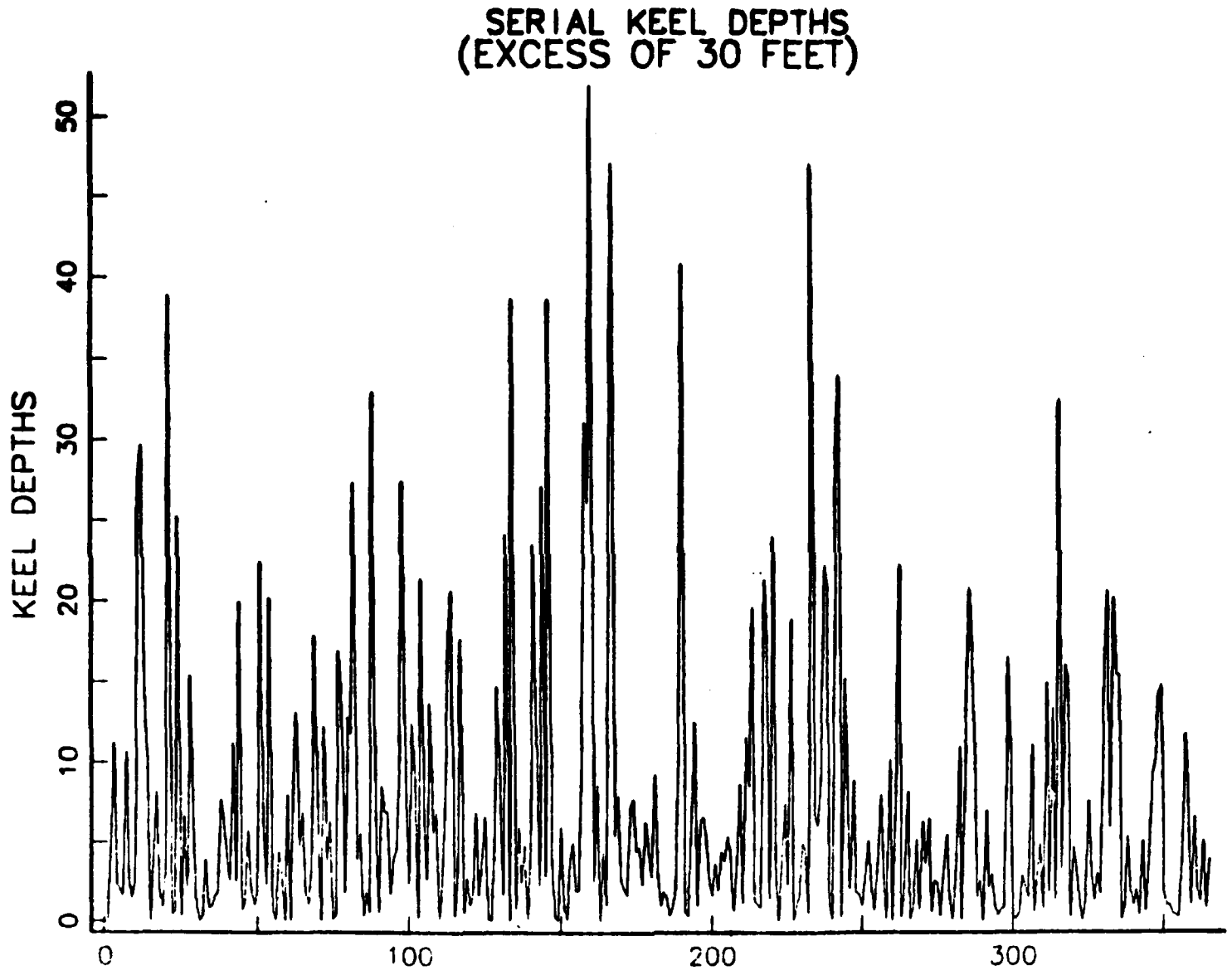

SERIAL KEEL DEPTHS
(EXCESS OF 30 FEET)
KEEL DEPTHS
0
10
20
30
40
50
0
100
200
300

FIGURE 20

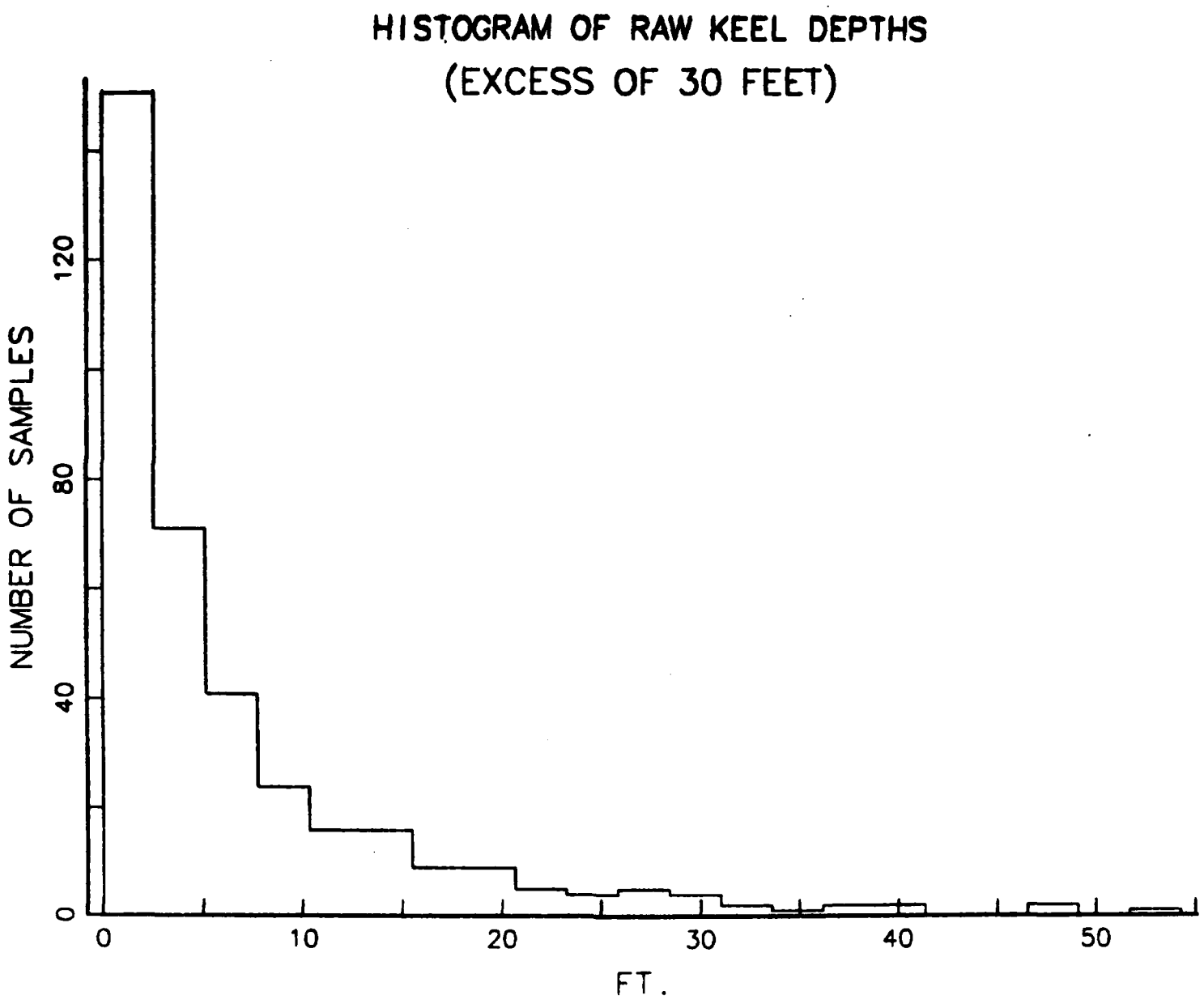

HISTOGRAM OF RAW KEEL DEPTHS
(EXCESS OF 30 FEET)
NUMBER OF SAMPLES
0
40
80
120
0
10
20
30
40
50
FT.

FIGURE 21

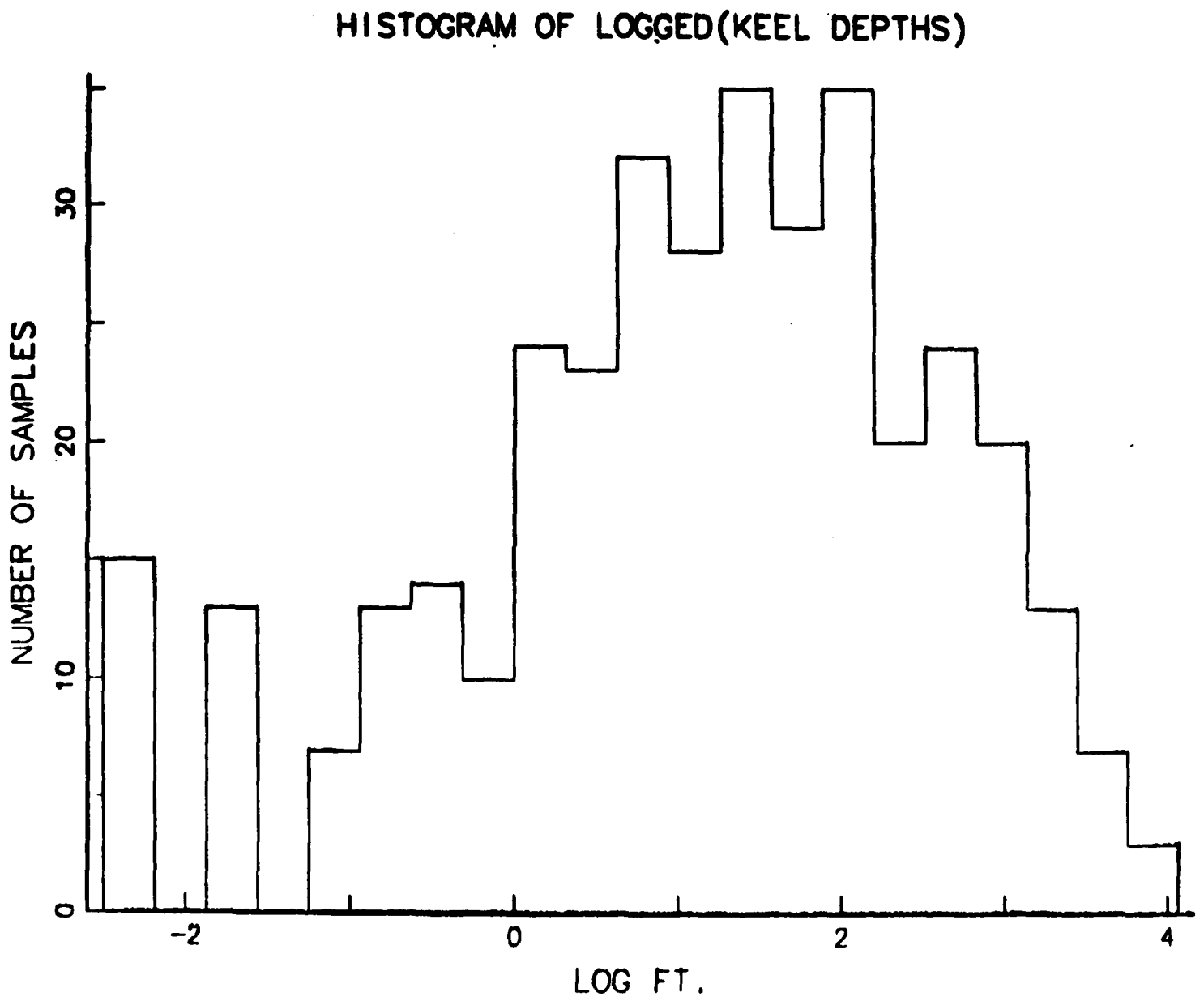

HISTOGRAM OF LOGGED(KEEL DEPTHS)
NUMBER OF SAMPLES
0
10
20
30
-2
0
2
4
LOG FT.

FIGURE 22

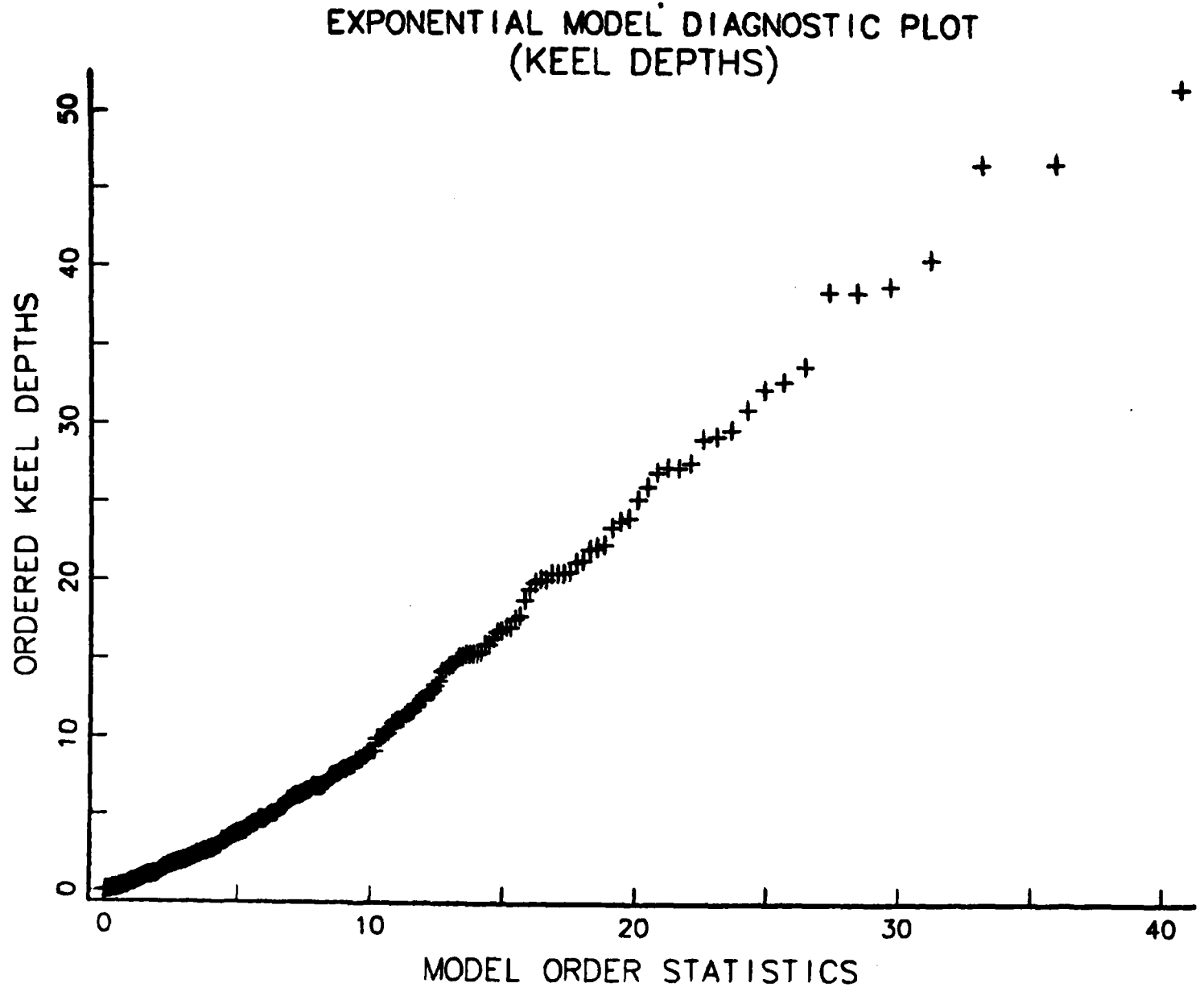

EXPONENTIAL MODEL DIAGNOSTIC PLOT
(KEEL DEPTHS)
ORDERED KEEL DEPTHS
0
10
20
30
40
50
0
10
20
30
40
MODEL ORDER STATISTICS

FIGURE 23

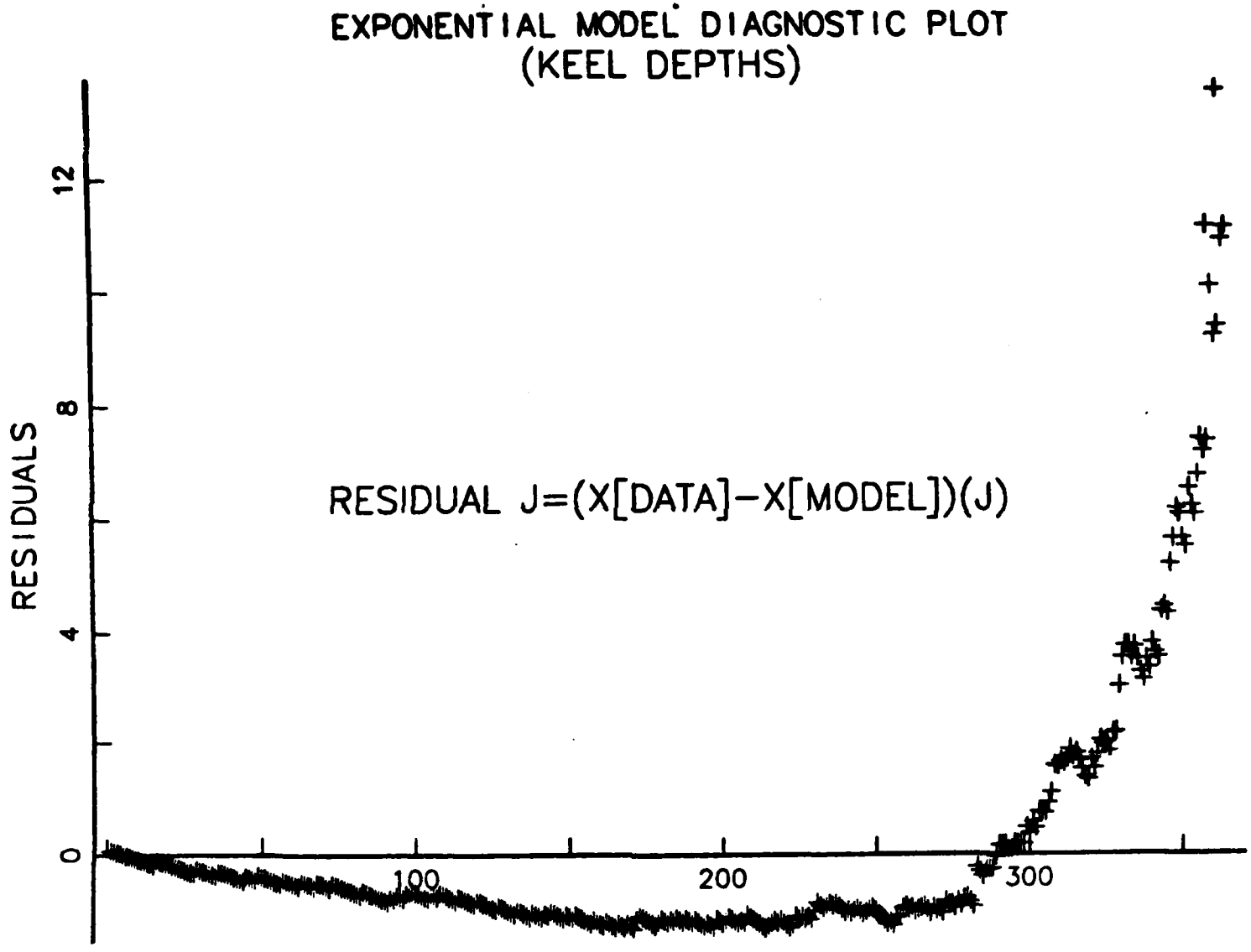

EXPONENTIAL MODEL DIAGNOSTIC PLOT
(KEEL DEPTHS)
RESIDUAL J=(X[DATA]-X[MODEL])(J)
RESIDUALS
0
4
8
12
100
200
300

FIGURE 24

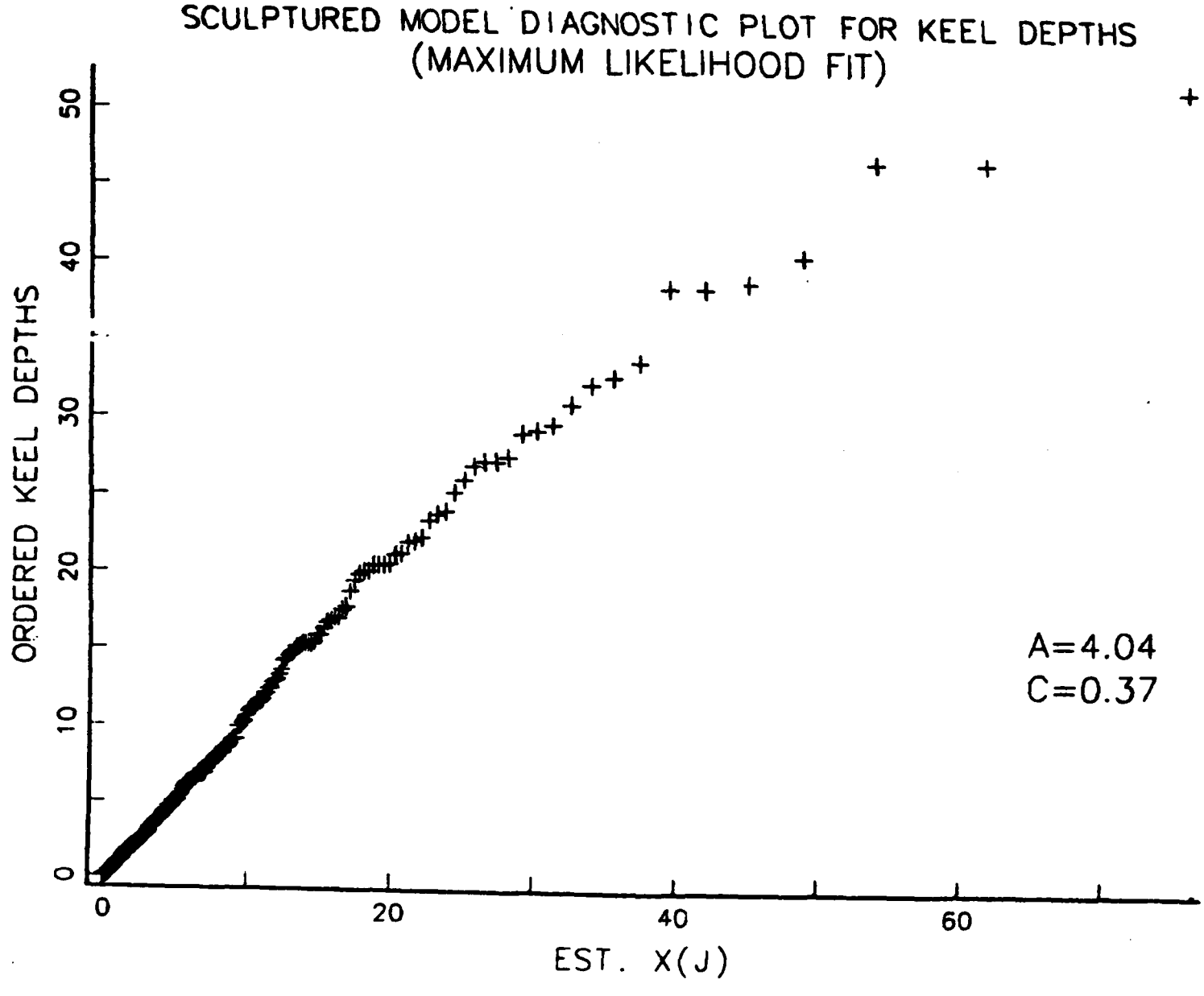

SCULPTURED MODEL DIAGNOSTIC PLOT FOR KEEL DEPTHS
(MAXIMUM LIKELIHOOD FIT)
ORDERED KEEL DEPTHS
EST. X(J)
0
10
20
30
40
50
0
20
40
60
A=4.04
C=0.37

FIGURE 25

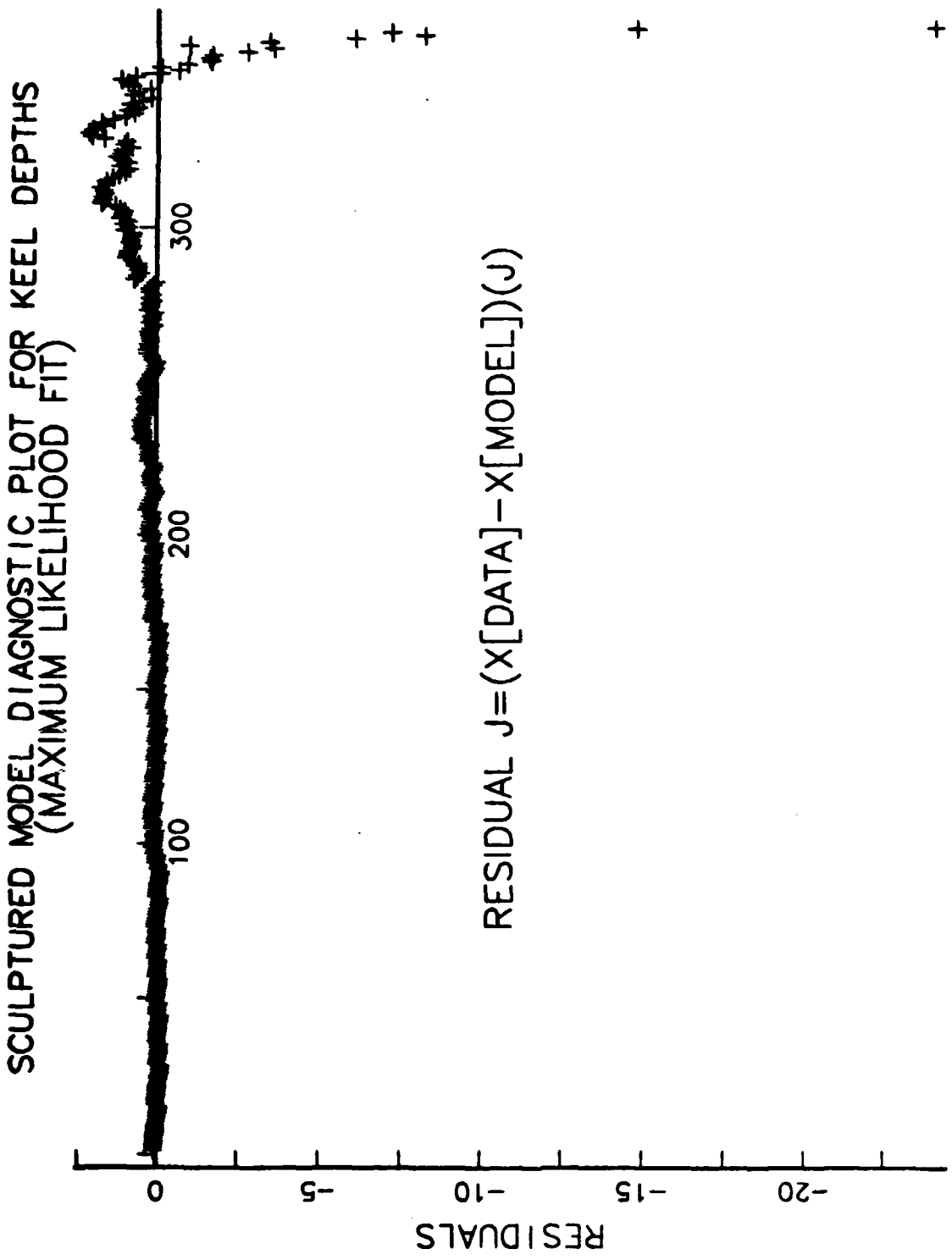

SCULPTURED MODEL DIAGNOSTIC PLOT FOR KEEL DEPTHS
(MAXIMUM LIKELIHOOD FIT)
RESIDUAL J=(X[DATA]-X[MODEL])(J)
RESIDUALS
0
-5
-10
-15
-20
100
200
300

FIGURE 26

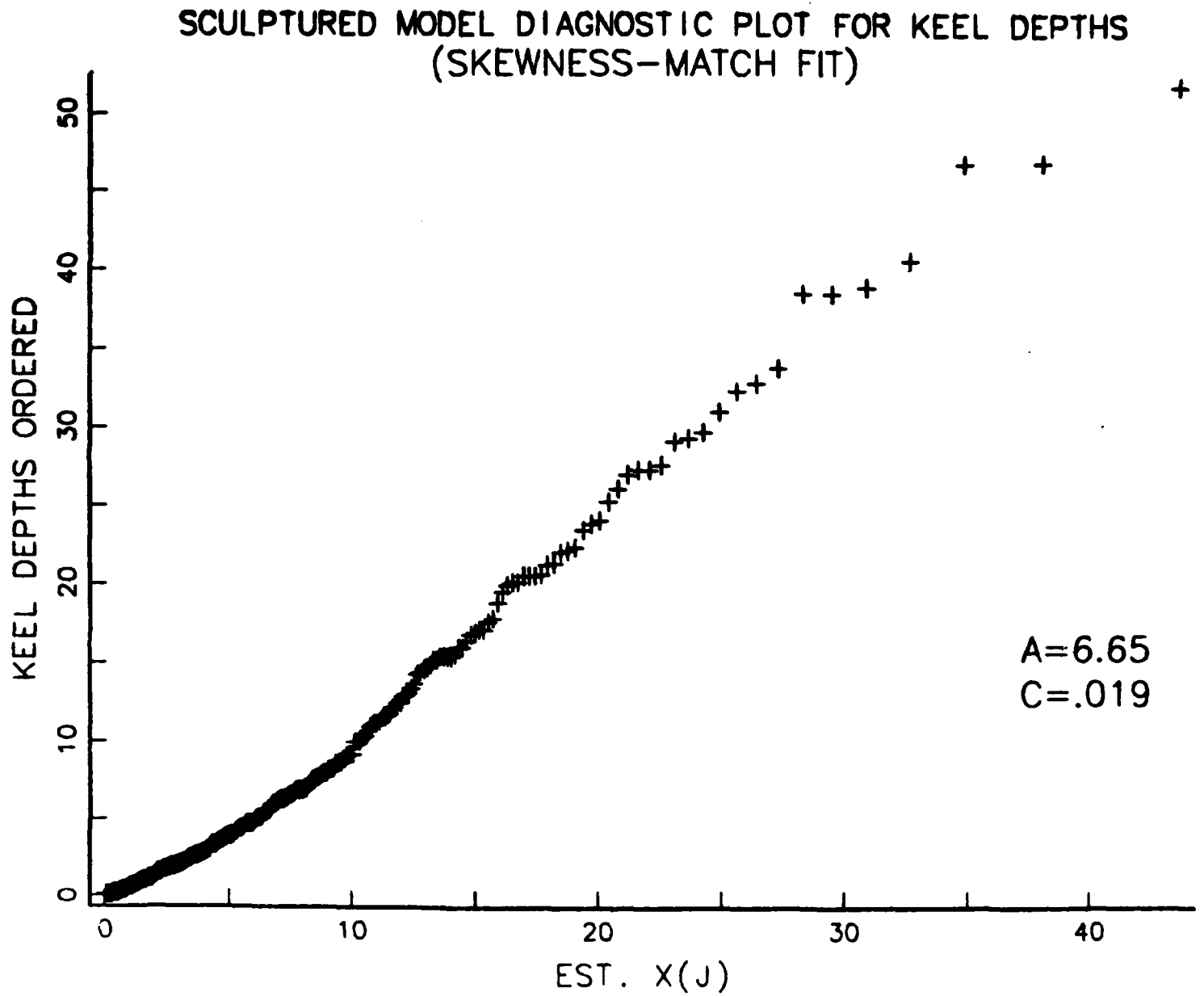

SCULPTURED MODEL DIAGNOSTIC PLOT FOR KEEL DEPTHS
(SKEWNESS-MATCH FIT)
KEEL DEPTHS ORDERED
EST. X(J)
0
10
20
30
40
50
A=6.65
C=.019

FIGURE 27

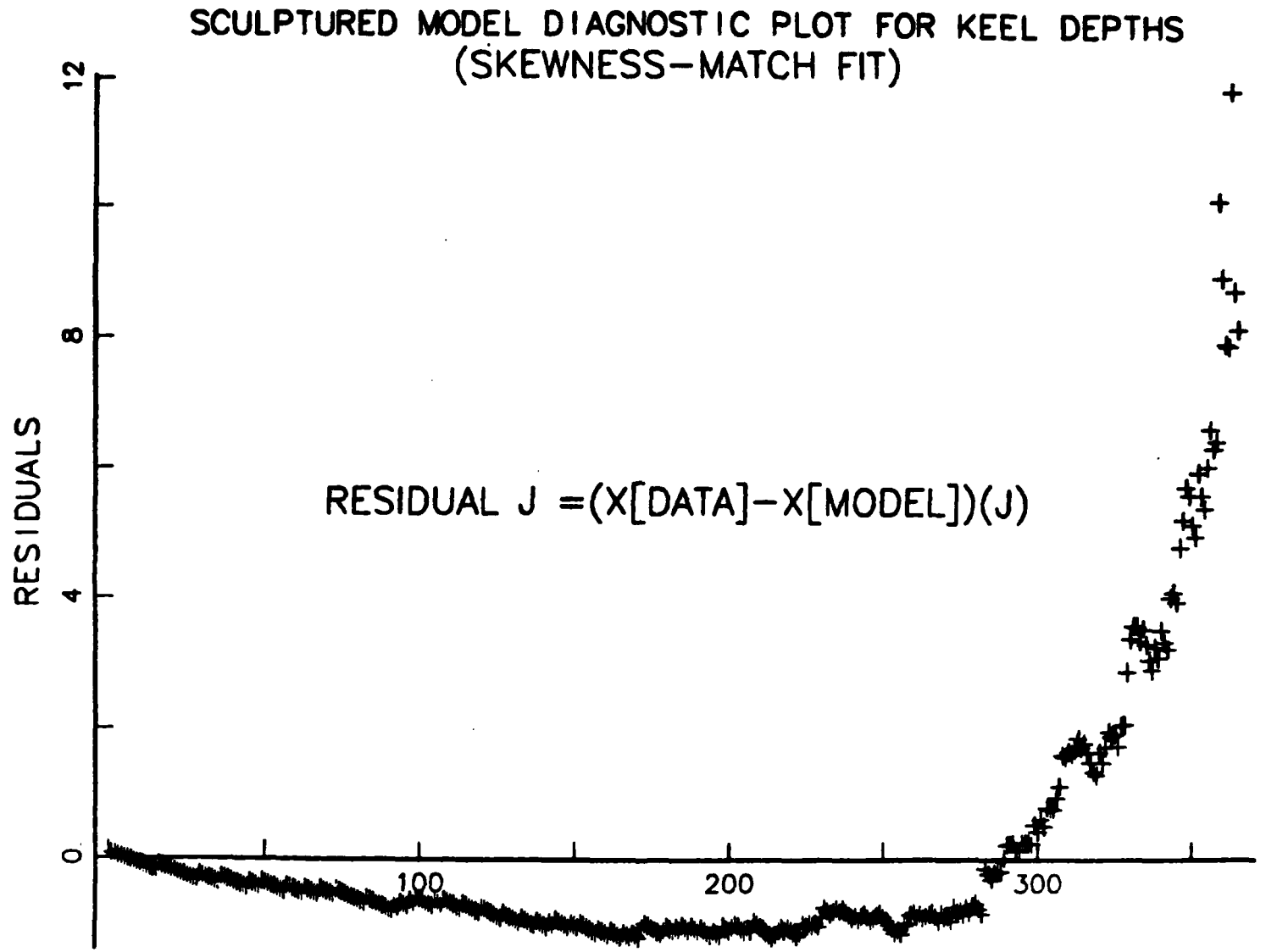

SCULPTURED MODEL DIAGNOSTIC PLOT FOR KEEL DEPTHS
(SKEWNESS-MATCH FIT)
RESIDUALS
12
8
4
0
RESIDUAL J =(X[DATA]-X[MODEL])(J)
100
200
300

FIGURE 28

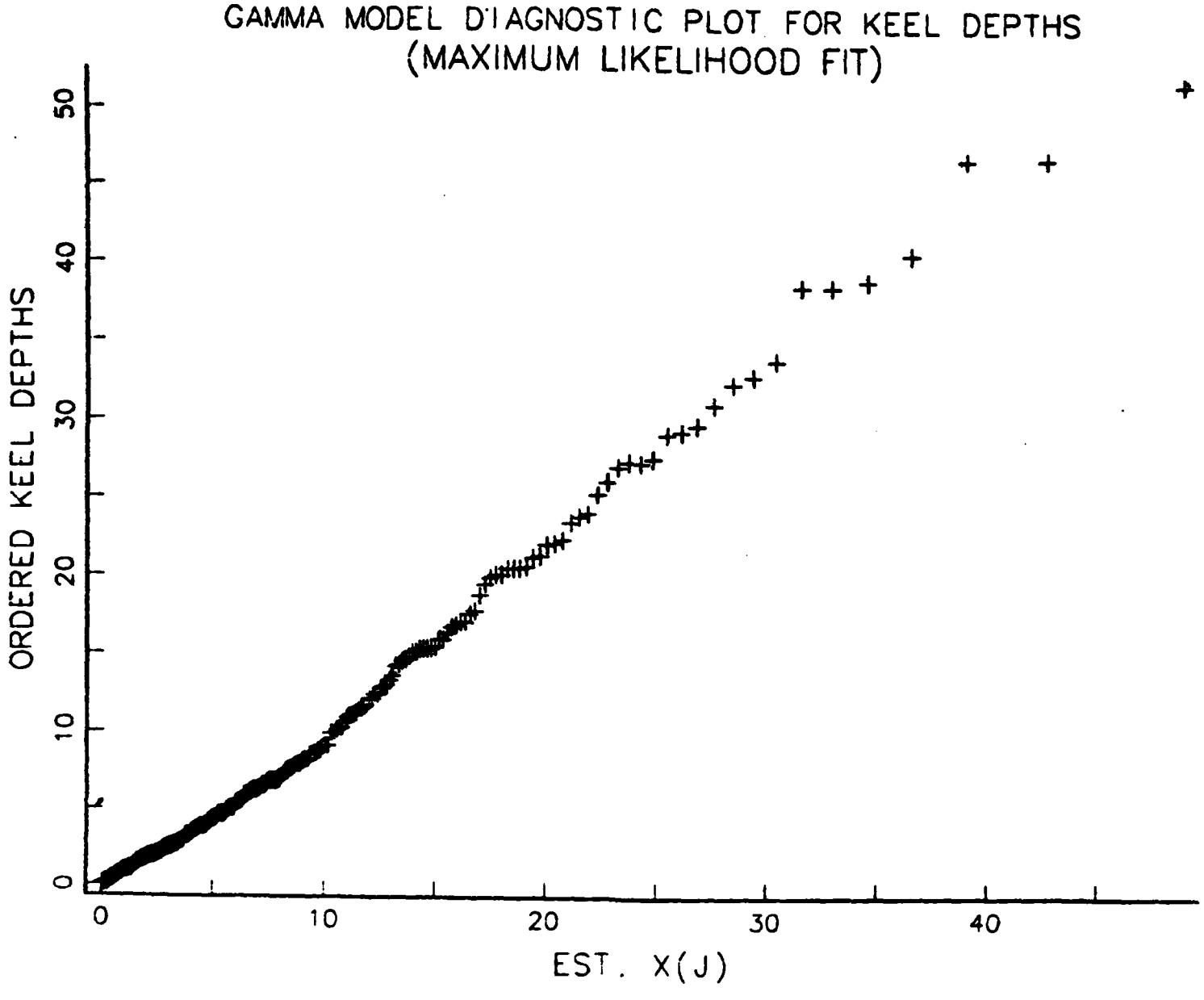

GAMMA MODEL DIAGNOSTIC PLOT FOR KEEL DEPTHS
(MAXIMUM LIKELIHOOD FIT)
ORDERED KEEL DEPTHS
0
10
20
30
40
50
0
10
20
30
40
EST. X(J)

FIGURE 29

GAMMA MODEL DIAGNOSTIC PLOT FOR KEEL DEPTHS
(MAXIMUM LIKELIHOOD FIT)

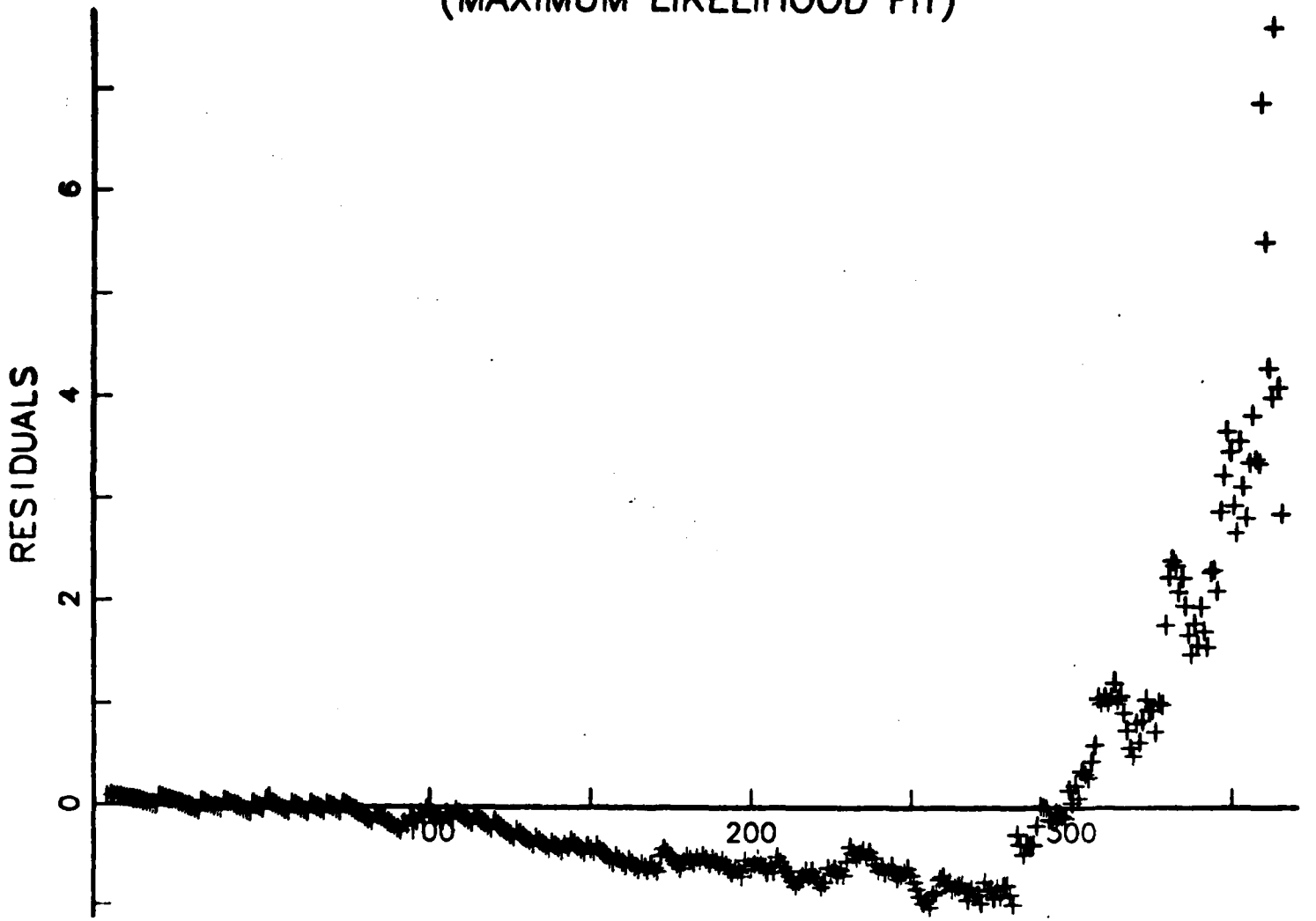


FIGURE 30

The exponentially sculptured model (6.14) was also fitted to the keel data by choosing C to match the coefficient of variation of the data and choosing A to match the mean. The values of the estimated parameters and the predicted moments are as follows.

Table 7

Coefficient of Variation - Match fit of Keel Depths by the Exponentially Sculptured Model

Estimates	Raw Data	Model
$\hat{E}[X]$	6.90	6.90
$\hat{C.V.}[X]$	1.26	1.26
$\hat{Skew}[X]$	2.22	4.01
$\hat{Kurt.}[X]$	5.55	37.86
Lower Quartile, $\underline{Q}$	1.20	1.64
Median	3.60	4.13
Upper Quartile, $\overline{Q}$	8.60	8.87

Diagnostic Plots appear as Figures 31 and 32.

As was found to be true for spacings, the maximum likelihood fitted sculptured model (6.1) applied to the present keel depth (reference d = 30 ft.) data systematically overestimates the magnitude of the far right tail (the number of deep keels). In this case the gamma model underestimates the final right tail values, see Figures 29 and especially 30. The skewness-matched sculptured exponential also tends to underestimate the far right tail of the data. An examination of the residual plot of Figure 26 suggests that the m.l.e.- fitted sculptured model nicely fits all keel size data except the very largest. The coefficient of variation match fit of exponentially sculptured model (6.14) produces residuals that are less structured

than either the gamma model or sculptured model (6.1), see Figure 33. The model (6.14) was also fit by matching the skewness, this fit produced residuals similar in appearance to Figure 28.

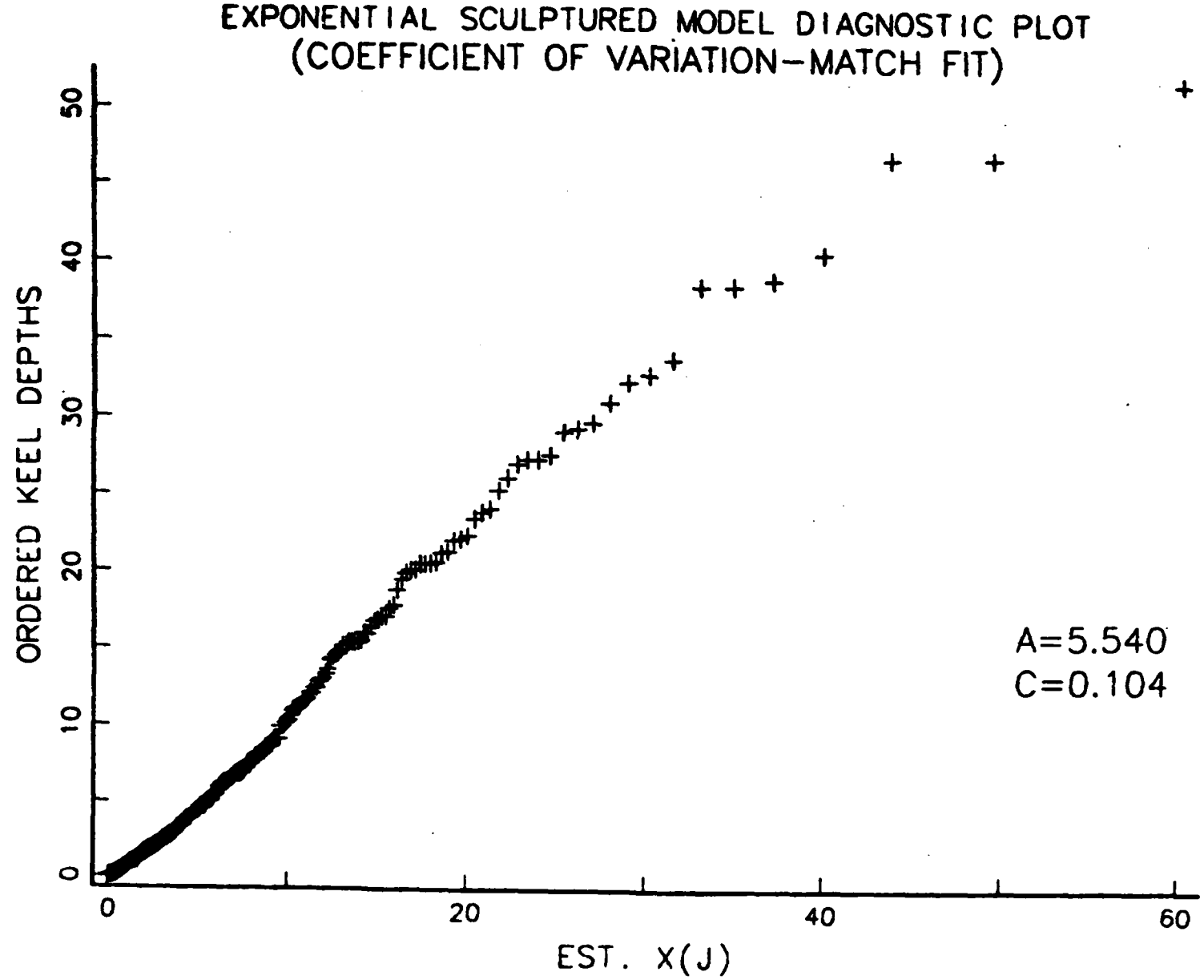

EXPONENTIAL SCULPTURED MODEL DIAGNOSTIC PLOT
(COEFFICIENT OF VARIATION–MATCH FIT)
ORDERED KEEL DEPTHS
0
10
20
30
40
50
0
20
40
60
EST. X(J)
A=5.540
C=0.104

FIGURE 31

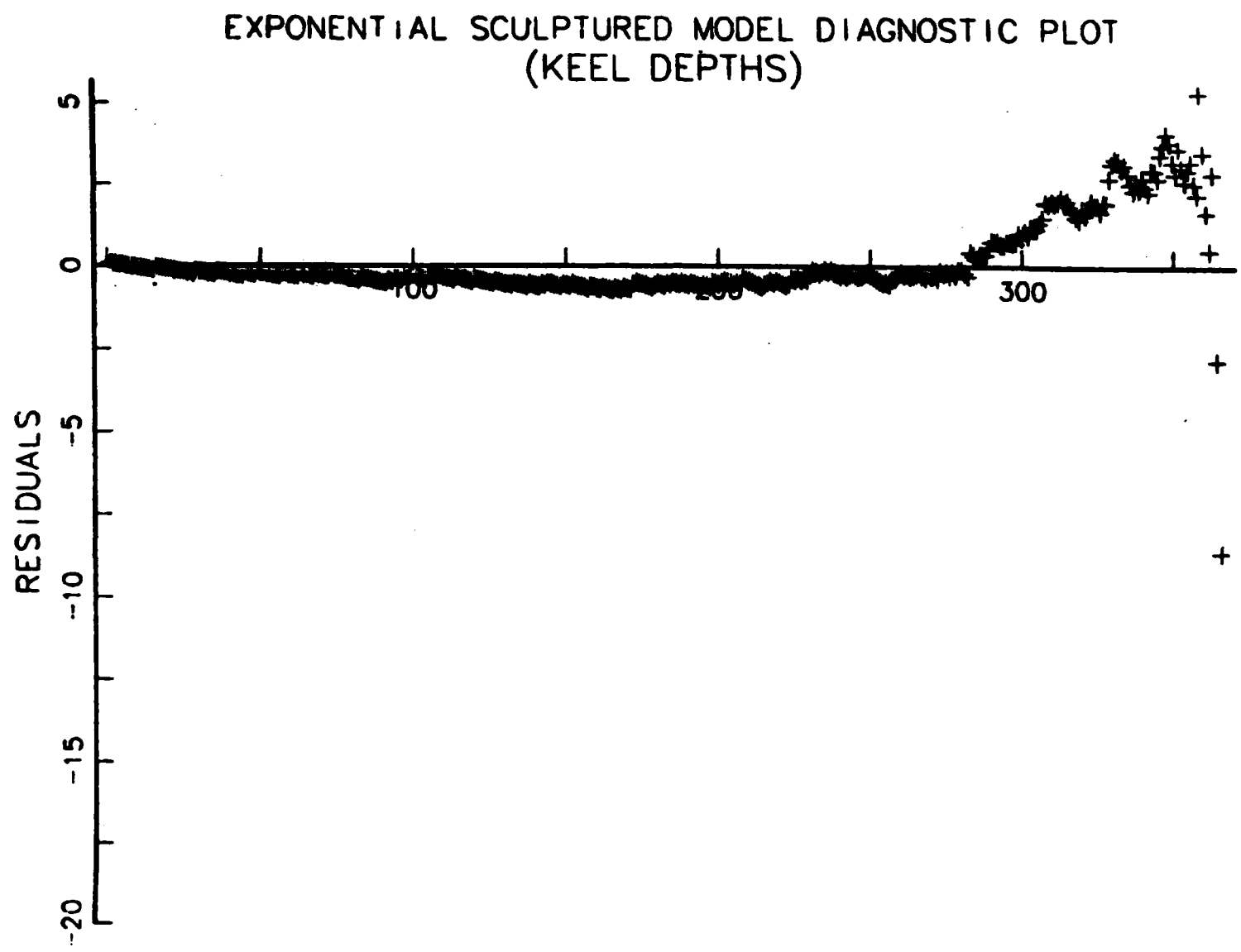

EXPONENTIAL SCULPTURED MODEL DIAGNOSTIC PLOT
(KEEL DEPTHS)
RESIDUALS
5
0
-5
-10
-15
-20
100
300

FIGURE 32

10. Acknowledgments

The authors wish to thank L. Uribe for his programming assistance. Figures 2-32 were produced by an experimental APL package GRAFSTAT2 which the Naval Postgraduate School is using under a test agreement with IBM Watson Research Center, Yorktown Heights, NY. We are grateful to Professor P. A. W. Lewis of the Naval Postgraduate School, and to Dr. P. D. Welch and Dr. P. Heidelberger of IBM for making this excellent data-analytic implement available to us.

References

1. Cramér, H. and Leadbetter, M. R. (1967). Stationary and Related Stochastic Processes. John Wiley and Sons, Inc. New York, NY.

2. Gaver, D. P. and Acar, M. (1979). Analytical hazard representations for use in reliability, mortality, and simulation studies. Commun. Statist.-Simula. Computat., B 8(2), pp. 91-111.

3. Gaver, D. P. and Jacobs, P. A. (1981). On combinations of random loads. S.I.A.M. J. Applied Math. Vol. 40, pp. 454-466.

4. Gaver, D. P. (1982). Sculptured distributions and some applications. Report in preparation.

5. Gaver, D. P. and Jacobs, P. A. (1982). Data Analysis and Modeling of Arctic Sea Ice Subsurface Roughness. Naval Postgraduate School Tech. Report.

6. Hibler, W. D. III, Weeks, W. F. and Moch, S. J. (1972). Statistical aspects of sea-ice ridge distributions. J. of Geophysical Research, Vol. 77, pp. 5954-5970.

7. McNeil, D. R. (1977). Interactive Data Analysis. Wiley-Interscience, New York, NY.

8. Shapiro, S. S. and Wilk, M. B. (1972). An analysis of variance test for the exponential distribution (complete samples). TECHNOMETRICS, Vol. 14, pp. 355-370.

9. Stephens, M. A. (1978). On the W test for exponentiality with origin known. TECHNOMETRICS, Vol. 20, pp. 33-36.

10. Tukey, J. W. (1977). Exploratory Data Analysis. Addison-Wesley Pub. Co., Reading, MA.

11. Wadhams, P. and Horne, R. J. (1980). An analysis of ice profiles obtained by submarine sonar in the Beaufort Sea. J. of Glaciology, Vol. 25, pp. 401-424.

12. Wilk, M. B. and Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. BIOMETRIKA, Vol. 55, pp. 1-17.

References

1. Cramér, H. and Leadbetter, M. R. (1967). Stationary and Related Stochastic Processes. John Wiley and Sons, Inc. New York, NY.

2. Gaver, D. P. and Acar, M. (1979). Analytical hazard representations for use in reliability, mortality, and simulation studies. Commun. Statist.-Simula. Computat., B 8(2), pp. 91-111.

3. Gaver, D. P. and Jacobs, P. A. (1981). On combinations of random loads. S.I.A.M. J. Applied Math. Vol. 40, pp. 454-466.

4. Gaver, D. P. (1982). Sculptured distributions and some applications. Report in preparation.

5. Hibler, W. D. III, Weeks, W. F. and Moch, S. J. (1972). Statistical aspects of sea-ice ridge distributions. J. of Geophysical Research, Vol. 77, pp. 5954-5970.

6. McNeil, D. R. (1977). Interactive Data Analysis. Wiley-Interscience, New York, NY.

7. Shapiro, S. S. and Wilk, M. B. (1972). An analysis of variance test for the exponential distribution (complete samples). TECHNOMETRICS, Vol. 14, pp. 355-370.

8. Stephens, M. A. (1978). On the W test for exponentiality with origin known. TECHNOMETRICS, Vol. 20, pp. 33-36.

9. Tukey, J. W. (1977). Exploratory Data Analysis. Addison-Wesley Pub. Co., Reading, MA.

10. Wadhams, P. and Horne, R. J. (1980). An analysis of ice profiles obtained by submarine sonar in the Beaufort Sea. J. of Glaciology, Vol. 25, pp. 401-424.

11. Wilk, M. B. and Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. BIOMETRIKA, Vol. 55, pp. 1-17.

APPENDIX

Estimates for the Percent-Points of Maximum Keel Depths, and Corresponding Uncertainty Estimates (Confidence Limits).

In Section 9 the exponentially sculptured model

$$X = AZe^{CZ} \tag{A-1}$$

was fitted to a batch of 365 keel depths; this is Model 3, (6.14). Furthermore, diagnostic plots of order statistics residuals indicated a reasonably successful fit of the data by the model. We now wish to utilize the model to predict statistical aspects of the maximum keel depth to be encountered in a further series of keel observations. Specifically we illustrate the procedures and results by assuming that

(a) a future sequence of 365 keel depths is of interest, and that these data come independently and randomly from model (A-1), or (6.14);

(b) we are interested in predicting the $95^{\underline{th}}$ percent point of the maximum of the data values in (a);

(c) we are also interested in associating 95% confidence limits with the point estimate of (b).

The above numerical values are illustrative only; it will be equally possible to predict the median or mean of the future maximum, together with confidence limits.

The form of the model (A-1) is especially convenient for addressing (b) and (c). Clearly the $95^{\underline{th}}$ percent point for the maximum of a (future) sample of 365 unit exponential random variables is

$$z_{(365)}(0.95) = -\left[\ell n\left\{1 - (0.95)^{1/(365)}\right\}\right] = 8.87\ldots ; \tag{A-2}$$

such numbers can be obtained accurately and handily from any set of extensive tables, or even from a hand-held calculator. Now by monotonicity, see (6.16), the $95^{\underline{th}}$ percent point of the maximum of a sample of 365 future keels is given by

$$x_{(365)}(0.95) = A\left[z_{(365)}(0.95)\right]e^{C[z_{(365)}(0.95)]} \qquad \text{(A-3)}$$

provided the model is correct and A and C are known. If the model is correct but $\hat{A}$ and $\hat{C}$ are estimates of A and C, then an estimate of the percent point is

$$\hat{x}_{(365)}(0.95) = \hat{A}\left[z_{(365)}(0.95)\right]e^{\hat{C}[z_{(365)}(0.95)]} \; . \qquad \text{(A-4)}$$

Thus a point estimate of the percent point of the maximum can be generated by simply substituting the parameter point estimates into (A-3), a very simple and direct task.

Next address (c), the uncertainty in the above estimate, or, more specifically, approximate confidence limits for the unknown percent point. We compute two estimates: the <u>jackknife</u> confidence limits, and the <u>bootstrap</u> confidence limits; see Efron [1980] for a leisurely discussion of both methods.

The <u>jackknife procedure</u> involves deletion of one observation at a time from the batch of data, and the re-computation of $\hat{A}$ and $\hat{C}$ using the remainder of the data; the estimates obtained omitting data point i are called $\hat{A}_{(i)}$ and $\hat{C}_{(i)}$. Next one computes the quantities

$$\hat{L}_{(i)} = \ell n\left\{\hat{A}_{(i)}\left[z_{(365)}(0.95)\right]e^{\hat{C}_{(i)}[z_{(365)}(0.95)]}\right\} ; \qquad \text{(A-5)}$$

the logarithm is taken in order to render jackknifing more valid by symmetrizing the sampling distribution of $\hat{x}_{(365)}(0.95)$. Finally one computes

$$\hat{L} = 365\,\hat{L}_{(0)} - \frac{364}{365}\sum_{i=1}^{365}\hat{L}_{(i)} , \qquad \text{(A-6)}$$

where $\hat{L}_{(0)} \equiv \ell n\,\hat{x}_{(365)}(0.95)$, the logged percent point estimate with no observations removed. The jackknifed estimate of the variance of $\hat{L}$ is here

$$\hat{VAR}[\hat{L}] = \frac{364}{365}\sum_{i=1}^{365}(L_{(i)} - \hat{L})^2 ; \qquad \text{(A-7)}$$

in the present instance the numerical value is 0.0156. The approximate 95% confidence interval for $\ell n\,x_{(365)}(0.95)$ is then

$$\hat{L} \pm 1.96\sqrt{\hat{VAR}[\hat{L}]} = [4.585,\ 5.075] ; \qquad \text{(A-8)}$$

the 95% confidence interval for the actual percent point of the maximum is then obtained by exponentiating the limits of (A-8), giving the interval [98.0, 160.0].

The <u>bootstrap procedure</u> involves re-sampling: from the empirical distribution of the original data points, obtain 200 independent random "bootstrap" samples of size 365 each, and from each bootstrap sample compute estimates of parameters A and C , and thence of $\hat{x}_{(365)}(0.95)$; there will thus be 200 bootstrap estimates in all. Now simply order the latter and find the $5^{\underline{th}}$ - smallest and largest ($195^{\underline{th}}$ - smallest) of these bootstrap

estimates, which will be quoted as the 95% confidence interval; and is [91.8, 149.9]. Notice that the log transformation is not required when the bootstrap method is utilized. It is gratifying that the bootstrap and jackknife methods are in such close numerical agreement for the present data. Both methods are somewhat more computationally demanding than, say, the classical "delta method" would be, but are entirely feasible using modern computers.

Reference

Efron, B. (1980). "The jackknife, the bootstrap, and other resampling plans". Tech. Report No. 63, Div. of Biostatistics, Stanford Univ., Stanford, CA.

DISTRIBUTION LIST

	NO. OF COPIES
Library, Code 0142 Naval Postgraduate School Monterey, CA 93940	4
Dean of Research Code 012A Naval Postgraduate School Monterey, CA 93940	1
Library, Code 55 Naval Postgraduate School Monterey, CA 93940	1
Professor D. P. Gaver Code 55Gv Naval Postgraduate School Monterey, CA 93940	155

END

FILMED

10-83

DTIC